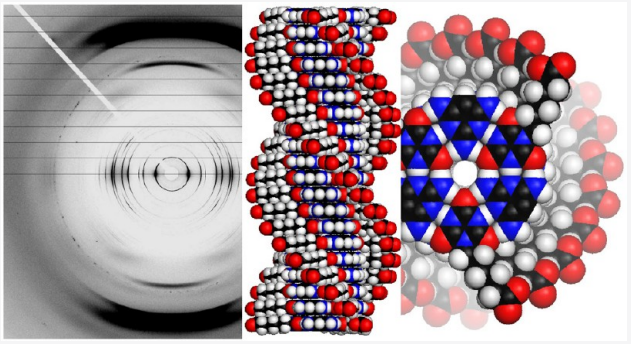
JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
pubs.acs.org/JACS
Article

X-ray Fiber Diffraction and Computational Analyses of Stacked Hexads in Supramolecular Polymers: Insight into Self-Assembly in Water by Prospective Prebiotic Nucleobases

Asem Alenaizan, Carlos H. Borca, Suneesh C. Karunakaran, Amy K. Kendall, Gerald Stubbs, Gary B. Schuster, C. David Sherrill,* and Nicholas V. Hud*

Cite This: J. Am. Chem. Soc. 2021, 143, 6079–6094

ABSTRACT: Aqueous solutions of equimolar mixtures of 2,4,6-triaminopyrimidine (TAP) and carboxylic acid substituted cyanuric acid (CyCo6 or *R*-4MeCyCo6) monomers self-assemble into gel-forming supramolecular polymers. Macroscopic fibers drawn from these mixtures were analyzed by X-ray diffraction to determine their molecular structures. Computational methods were used to explore the intrinsic intermolecular interactions that contribute to the structure and stability of these assemblies. Both polymers are formed by the stacking of hexameric rosettes, $(TAP/CyCo6)_3$ or $(TAP/R\text{-}4MeCyCo6)_3$, respectively, into long, stiff, twisted stacks of essentially planar rosettes. Chiral, left-handed supramolecular polymers with a helical twist angle of −26.7° per hexad are formed when the pure enantiomer *R*-4MeCyCo6 is used. These hexad stacks pack into bundles with a hexagonal crystalline lattice organization perpendicular to the axis of the macroscopic fiber. Polymers formed from TAP and CyCo6, both of which are achiral, assemble into macroscopic domains that are packed as a centered rectangular lattice. Within these domains, the individual polymers exist as either right-handed or left-handed helical stacks, with twist angles of +15° or −15° per hexad, respectively. The remarkable ability of TAP and cyanuric acid derivatives to self-assemble in water, and the structural features of their supramolecular polymers reported here, provide additional support for the proposal that these heterocycles could have served as recognition units for an early form of nucleic acids, before the emergence of RNA.

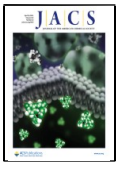


INTRODUCTION

Beginning in 1990, the Lehn and Whitesides laboratories independently reported using triazines and pyrimidines, and their modified forms, to create complex structures in nonpolar solvents.[1,2] While mixtures of one triamino species, melamine (MA) or 2,4,6-triaminopyrimidine (TAP), with one triketo species, cyanuric acid (Cy) or barbituric acid (BA), result in the formation of unbounded sheet-like structures of pairing monomers,[3] modifying one hydrogen-bonding "face" of one pairing species blocks the formation of sheet-like assemblies while maintaining hydrogen-bonded structures. In particular, the two unmodified faces of these molecules enable the formation of assemblies named "rosette", "linear tape", and "crinkled tape" (Figure 1).[1,2,4−16] The two tape assemblies remain unbounded in one dimension and therefore are potentially of infinite length, making them attractive for extended assemblies.[6,9,16] The hexad rosette structure, on the other hand, being restricted to six monomers in a specific planar arrangement,[17] is attractive for the rational design of supramolecular assemblies.[14−16,18−21] After decades of research on these pairing systems, the only reported crystal structure for stacked hexameric rosettes containing one of these triamino species paired with a triketo species is of two stacked MA−BA hexads with bulky substituents attached to each heterocycle. It is likely that the bulky substituents prevent extended stacking of hexads and cause the observed deformation of the hexad planar geometry.[5]

We anticipated that hexad rosettes would stack in water to form linear assemblies if one of the two hexad-forming heterocycles was modified with a solubilizing, electrostatically charged group. Indeed, TAP modified with an amide-linked succinic acid moiety (TAPAS, Figure 1D) and Cy produce hydrogel-forming supramolecular polymers consistent with a stacked-hexad assembly (Figure 2B).[22] Subsequent studies of TAP mixed with Cy modified with achiral (CyCo6) and chiral (*R*-4MeCyCo6) versions of a hexanoic acid tail (Figure 1D)

Received: November 18, 2020
Published: April 14, 2021



https://doi.org/10.1021/jacs.0c12010
J. Am. Chem. Soc. 2021, 143, 6079−6094

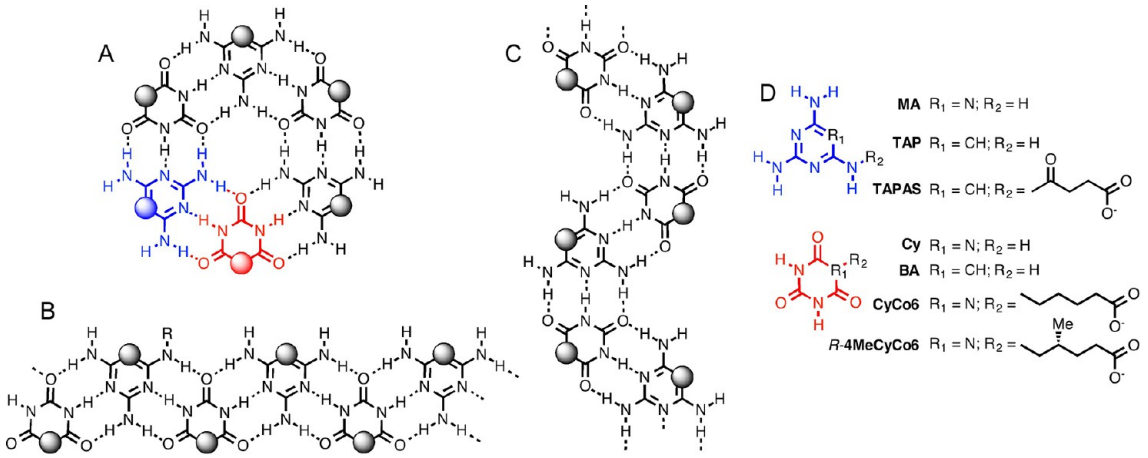


Figure 1. Hydrogen-bonded structures formed by a triamino monomer, melamine (MA) or 2,4,6-triaminopyrimidine (TAP) (blue structure), and a triketo monomer, barbituric acid (BA) or cyanuric acid (Cy) (red structure). In their unmodified forms, mixtures of one triamino heterocycle and one triketo heterocycle form insoluble, extended hydrogen-bonded assemblies. However, soluble assemblies result when one of the pairing "faces" of either molecule contains a modification (represented by a sphere) that blocks pairing with its complement. These monomers form assemblies designated (A) hexad rosette, (B) linear tape, and (C) crinkled tape. (D) Modified monomers discussed in this work that form water-soluble supramolecular assemblies.

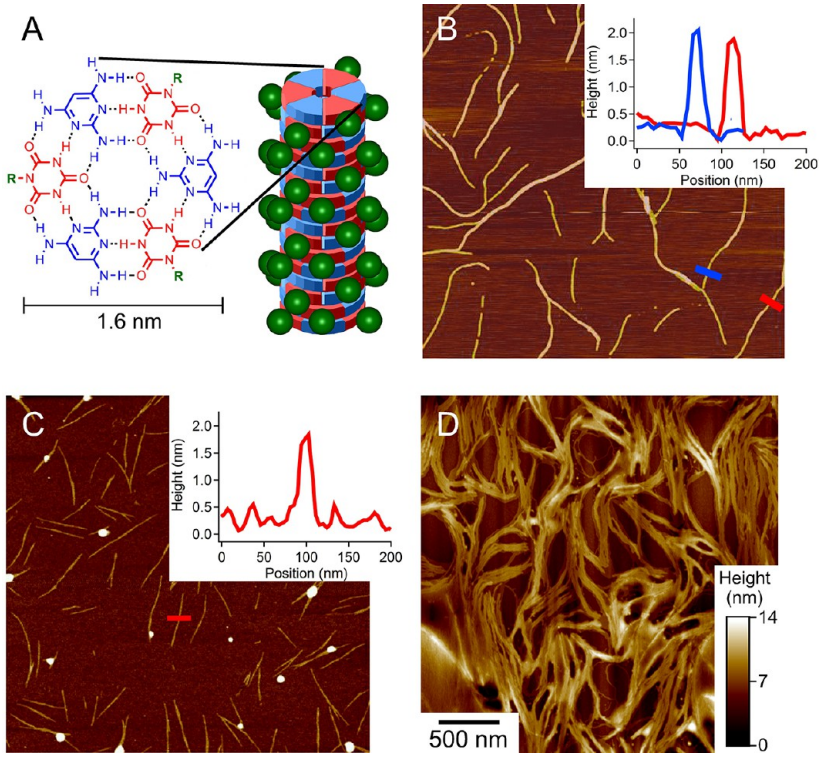


Figure 2. (A) Model of the supramolecular assembly for stacked hexads. (B) Atomic force microscopy (AFM) image of TAPAS−Cy assemblies, formed from 5 mM of each monomer. Reproduced from Cafferty et al., *J. Am. Chem. Soc.* **2013**, *135*, 2447−2450. Copyright 2013 American Chemical Society. (C) AFM image of TAP−CyCo6 assemblies, formed from 30 mM of each monomer in the presence of 1 mM of ethidium bromide (which was added to reduce polymer length and aggregation, as demonstrated in ref 23). (D) AFM image of TAP−*R*-4MeCyCo6 assemblies, formed from 15 mM of each monomer. The insets in panels B and C show height profiles measured where indicated by the red and blue lines drawn on the AFM images. The three AFM images are shown at the same magnification, as indicated by the 500 nm scale bar in D.

produced similar supramolecular polymers (Figure 2C,D).[23] Likewise, other modified anionic forms of TAP, BA, Cy, and MA were found to produce supramolecular polymers with the same morphology when mixed with an unmodified pairing partner.[24−26]

The assembled monomers that comprise supramolecular polymers in solution are bound by evanescent noncovalent interactions: typically, hydrogen bonds and, in aqueous solution, hydrophobic interactions. The relatively weak forces binding the monomers give rise to the characterization of these materials as "soft matter" and account for their ability to respond readily to external stimuli.[12,16,27−29] These characteristics also make detailed structural analysis of supramolecular polymers challenging.[30] Viscometric and light-scattering

https://doi.org/10.1021/jacs.0c12010
J. Am. Chem. Soc. 2021, 143, 6079−6094

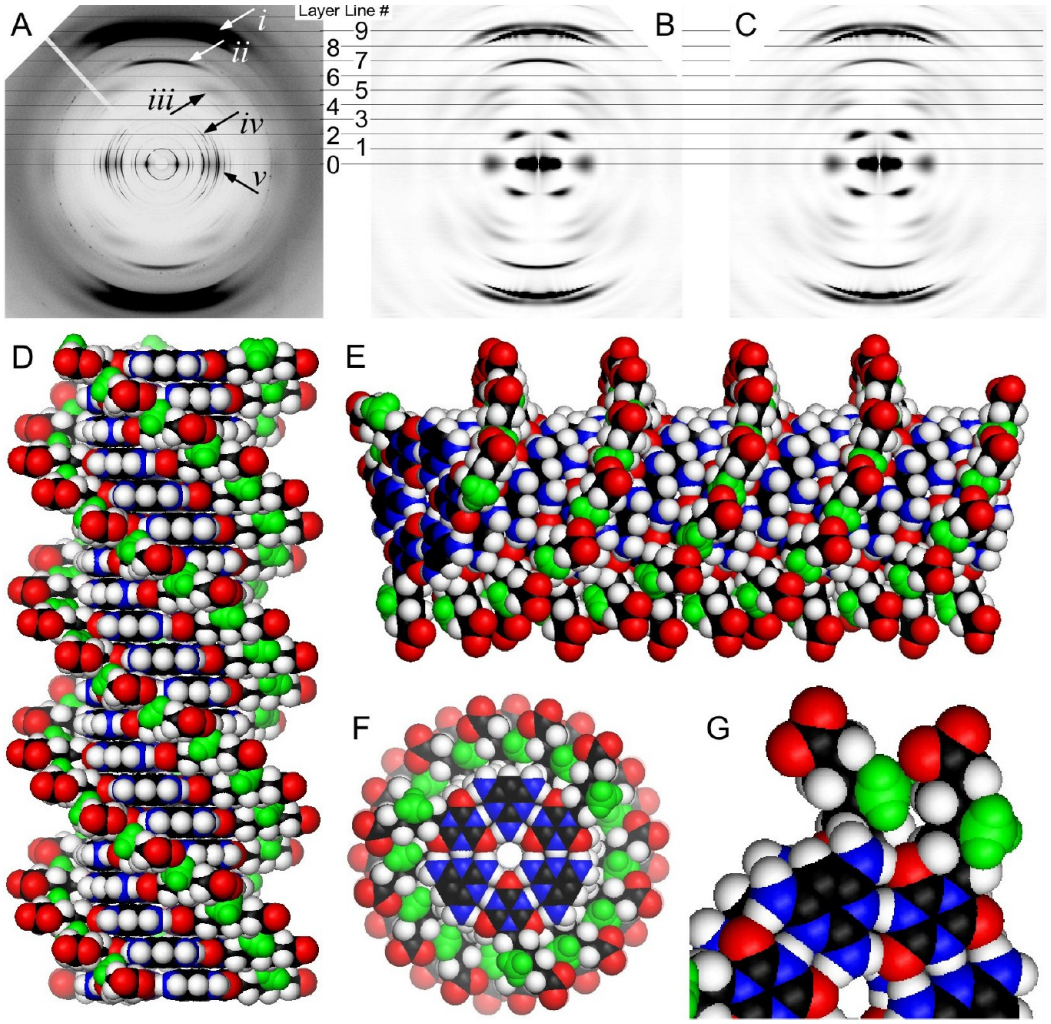


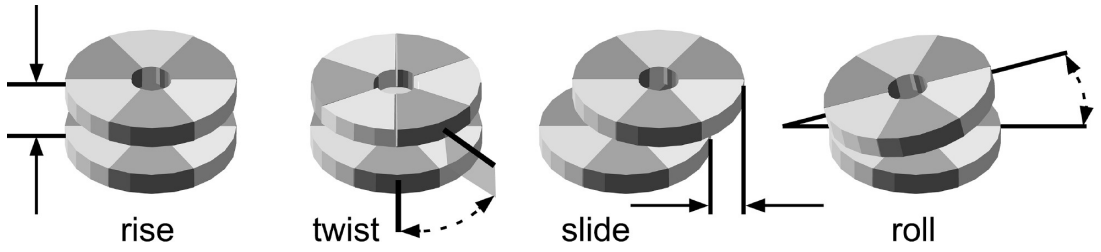
Figure 3. (A) X-ray diffraction pattern of a fiber composed of TAP and *R*-4MeCyCo6 in a 1:1 ratio. The orientation of the fiber is vertical relative to the diffraction pattern. The strong reflection, marked *i*, is at a reciprocal distance of 3.4 Å from the equator and indicative of hexads stacked along the fiber axis. Other features labeled *ii*–*v* are discussed in the text. (B) Calculated diffraction pattern based on a 100-hexad stack with a helical twist of −26.7° per hexad. The pattern shown was generated by the summing of individual patterns of the 100-hexad stack precessed about the *z*-axis with variable angles of tilt from the *z*-axis. Additional details on the generation of this pattern, including weighting of patterns with different hexad stack tilt angles, are provided in the text and in Figure 5. (C) Patterns generated in the same way as the pattern shown in B, except for hexad stacks with a helical twist of +26.7° per hexad. (D–F) Three views of an idealized model of TAP–*R*-4MeCyCo6 assemblies with a helical twist of −26.7° per hexad. (G) Close-up images of the methyl groups on the tails of two adjacent hexads from the model structure shown in D–F. The methyl groups of the *R*-4MeCyCo6 residues are shown in green.

measurements provide valuable information about the size and shape of supramolecular polymers but offer little insight into their molecular structures. The common spectroscopic tools typically do not provide a detailed view of the molecular organization of supramolecular polymers.[31] Optical, electron (EM), and force (AFM) microscopies provide useful images[32] but are incapable of the resolution required to reveal the critical details of the interactions between monomeric units. The application of X-ray crystallography to obtain an atomic level view of supramolecular polymers that form gels is challenging because removing the solvent (to form a xerogel or single crystal) can alter the polymer's structure in solution in unknown ways.[33] For example, Fenniri and co-workers prepared bicyclic self-pairing molecules that form supramolecular polymers of planar, stacked hexads in solution but crystallize as nonplanar, unstacked hexads.[34] In contrast, X-ray fiber diffraction (XRFD) provides a powerful tool for structural analysis that can preserve the linear, polymeric nature of supramolecular assemblies.[35−37] This technique has been successfully applied to supramolecular hydrogels for determining the unit cell dimensions and to confirm self-sorting in two-component supramolecular gels.[38−40]

The goal of this study is to uncover the structural features and physical principles that provide $(TAP/Cy)_3$ hexameric systems with unusually favorable properties, even among more highly engineered self-assembling molecules, including assembly at low millimolar concentrations in water, extraordinarily long persistence length, a strong propensity to spontaneously form macroscopic domains of uniform supramolecular chirality, and a hypersensitivity to the introduction of chiral monomers.[22,41] We report an XRFD study of $(TAP/CyCo6)_3$ and $(TAP/R\text{-}4MeCyCo6)_3$ polymers that provides the first direct support for the formation of stacked hexad assemblies with paired heterocycles as shown in Figure 1A. Previously, existence for these structures was inferred based on fiber dimensions measured by EM and AFM.[22] Additionally, fiber

https://doi.org/10.1021/jacs.0c12010
J. Am. Chem. Soc. 2021, 143, 6079−6094

Scheme 1. Definition of Helical Parameters for Stacked Hexads within Supramolecular Assemblies



diffraction patterns reveal that the helical twist of hexad stacks is extremely well-defined within the fiber state. Moreover, the adoption of different, but mutually defined, helical twists between fibers containing CyCo6 and those containing *R*-4MeCyCo6 demonstrates that adjacent hexads can adjust their twist angle to optimize substituent interactions and fiber lattice contacts.

The underpinnings of the hexad structure at different hierarchies of organization were explored using a complementary set of computational tools. Quantum mechanical (QM) computations were used to examine the intrinsic noncovalent interactions between two model hexad rosettes at various geometric configurations in the absence of complicating environmental effects (*e.g.*, solvation). The QM computations guided the choice of approximate classical methods for performing molecular dynamics (MD) simulations on larger hexad stacks. The MD simulations were used to investigate the structure of these assemblies in solution and to examine solvation effects on the stability of these systems.

Lastly, the combined results from the atomistic QM and MD simulations and experimental data were used to explain the observed macroscopic properties of these noncovalent assemblies. The QM and MD simulations show the importance of stacking interactions for the stability of these systems and support the conclusion that stacks of $(TAP/Cy)_3$ hexads have an unusually strong propensity to form helical structures.[41] The computational studies support the recent conjecture that a small cavity between stacked hexads might be able to trap ions,[42] similar to the central pore of G-tetrads.[43,44] The combined computational and experimental results provide insight into the macroscopic features exhibited by hexad assemblies and support the hypothesis that the unique self-assembly properties of BA, TAP, MA, and Cy, and their suspected prebiotic origin,[45] make them strong candidate recognition units for the informational polymers of life before the emergence of RNA.[24,46−49]

RESULTS

Fiber Structure of TAP−*R*-4MeCyCo6 Hexad Supramolecular Assemblies. Following the protocol described by Davies and co-workers for preparation of guanosine monophosphate (GMP) fibers,[50] we produced macroscopic fibers from 1:1 mixtures of TAP with CyCo6 and from 1:1 mixtures of TAP with *R*(+)-4-methyl-CyCo6 (*R*-4MeCyCo6), a chiral derivative of CyCo6 (Figure 1 D). All fibers examined by X-ray diffraction produced patterns that indicated high local order, but also signs of axial disorder. In particular, sharp arc intensities along the pattern's equator indicate crystalline packing perpendicular to the fiber axis, while more diffuse intensities along the layer lines are typical of noncrystalline fiber diffraction patterns and consistent with these fibers lacking long-range rotational or translational registration along the fiber axis. Fortunately, like the X-ray diffraction patterns obtained for duplex DNA,[51,52] triplex DNA,[53] and G-quartet assemblies,[50] the diffraction patterns exhibited by TAP−CyCo6 and TAP−*R*-4MeCyCo6 fibers provide sufficient information to obtain precise measurements of the helical parameters essential for constructing meaningful model structures.

Hexad Spacing and Helical Twist. The most prominent feature of the TAP−*R*-4MeCyCo6 diffraction pattern is the very strong near-meridional intensity at the reciprocal distance from the equatorial line of 3.4 Å (Figure 3A, *i*). This feature is characteristic of heterocycles stacked along the fiber axis, like that observed in the fiber diffraction patterns of DNA.[50−53] For TAP−*R*-4MeCyCo6 fibers this strong intensity is consistent with the planes of the hexads being perpendicular to the fiber axis with an interhexad rise of 3.4 Å along the fiber axis (Scheme 1).

The diffracted intensities in this pattern were best fit by nine layer lines, with the 3.4 Å intensity on the ninth layer line. The arcs observed on the second layer line were particularly important for defining the number of layer lines most appropriate for this pattern (Figure 3A, *iv*). A fiber diffraction pattern best-fit with nine layer lines usually indicates a helical repeat of nine residues. That is, a helical twist (Scheme 1) of 360° per nine hexads or, equivalently, a helical twist of 40° per hexad. However, each TAP−*R*-4MeCyCo6 hexad has 3-fold symmetry (Figure 2A). Consequently, the nine layer lines of this pattern could indicate nine hexads per 120° of helical rotation (a helical twist of 13.3° per hexad) or nine hexads per 240° of helical rotation (a helical twist of 26.7° per hexad). Therefore, analysis of additional features of the diffraction pattern was necessary to determine the actual helical twist of the assemblies.

The TAP−*R*-4MeCyCo6 diffraction pattern exhibits an arc centered along the seventh layer line (Figure 3A, *ii*). Simulated diffraction patterns of a TAP−*R*-4MeCyCo6 hexad stack with a helical twist of 26.7° also exhibit a centered arc of intensity at this location (Figure 3B). The vertical position of this arc depends on the helical twist of the stack, progressively moving from the ninth layer line to the midpoint between the ninth layer line and the equator (or zeroth layer line) as the helical twist of the hexad stack is increased from 0° to 60° (Supporting Information (SI), Figure S1). Thus, the appearance of this arc along the seventh layer line provides strong support for assemblies within TAP−*R*-4MeCyCo6 fibers having a helical twist of 26.7° per hexad.

The intensities along the second layer line of the calculated pattern are also consistent with the arcs observed along the second layer line in the experimental diffraction pattern (Figure 3B). The vertical position of these intensities in the calculated patterns also changes as the helical twist of a model stack is varied from 0° to 60°, moving progressively from the

https://doi.org/10.1021/jacs.0c12010
J. Am. Chem. Soc. 2021, 143, 6079−6094

equator to a position coinciding with the upper arc at the midpoint between the ninth layer line and the equator (SI, Figure S1). Thus, the location of these intensities along the second layer line in the experimental diffraction pattern provides additional support for a helical twist of 26.7° per hexad. We note that the pattern shown in Figure 3B was calculated using *one* stack of 100 hexads, as opposed to multiple stacks in a fiber matrix, for both the practical reason of computational expense and to identify features of the experimental pattern that result from order within individual hexad stacks, as opposed to the crystalline matrix of the fiber. Accordingly, the broad intensities observed along the second layer line in the calculated pattern become arcs in patterns calculated with multiple hexad stacks arranged with regular spacings, as illustrated in the following section.

Another distinctive feature of the experimental TAP−*R*-4MeCyCo6 diffraction pattern is the diffuse intensities that lie along the fifth layer line (Figure 3A, *iii*). Similar diffuse intensities appear along the fifth layer line of the calculated diffraction pattern of TAP−*R*-4MeCyCo6 with a helical twist of −26.7° (Figure 3B). These diffuse intensities are associated with the 4-methylhexanoic acid tail of *R*-4MeCyCo6, as these features do not appear in calculated diffraction patterns of TAP−Cy hexad stacks (SI, Figure S2). A lighter diffuse intensity also appears along the fourth layer line of the calculated diffraction pattern with a helical twist of −26.7° (Figure 3B). Increasing the helical twist value of the TAP−*R*-4MeCyCo6 hexad stack by 3.3°, to a twist of 30° per hexad, results in the movement of these two sets of diffuse intensities to the same location at the midpoint between the fourth and fifth layer lines (SI, Figure S1). Thus, the separation and appearance of these intensities along the fourth and fifth layer lines serve as a very sensitive metric for determining the helical twist of the TAP−*R*-4MeCyCo6 assembly. We note that in the experimental diffraction pattern, shown in Figure 3A, arcs of intensities appear along the fourth layer line, not diffuse intensities. This difference is consistent with these intensities in the experimental pattern being broken into discrete arcs by the transform of the crystal packing matrix (discussed below).

Handedness. The diffuse intensities along the fifth layer line also reveal the helical handedness of the TAP−*R*-4MeCyCo6 assembly. Because *R*-4MeCyCo6 is a chiral molecule, the left-handed and the right-handed helices of stacked TAP−*R*-4MeCyCo6 hexads are diastereomeric isomers of each other, not mirror images. Moreover, the most significant structural difference between the left-handed and the right-handed helical forms of a TAP−*R*-4MeCyCo6 assembly will be at the single chiral center of *R*-4MeCyCo6, the methyl group of the 4-methylhexanoic acid tail. Therefore, as expected, the calculated diffraction pattern of stacked TAP−*R*-4MeCyCo6 hexads with a helical twist of −26.7° per hexad (a left-handed helix, Figure 3B) is different from the pattern calculated for the same stack with a helical twist of +26.7° per hexad (a right-handed helix, Figure 3C). Most apparent is the permutation of the diffuse intensities along the fourth and fifth layer lines. The correspondence of the greater intensity along the fifth layer line with the pattern calculated for stacked TAP−*R*-4MeCyCo6 hexads with a helical twist of −26.7° supports a left-handed helix.

Based on the information extracted from the fiber diffraction pattern of TAP−*R*-4MeCyCo6 assemblies, we present a model for an individual stack of TAP−*R*-4MeCyCo6 hexads having a helical rise of 3.4 Å along the fiber axis and a helical twist of −26.7° per hexad (Figure 3D−F). A striking feature that emerges from this structure is the placement of the methyl groups of the *R*-4MeCyCo6 residues. Each of these methyl groups is within van der Waals contact of the 4-methylhexanoic acid tail of an *R*-4MeCyCo6 residue of an adjacent hexad, which explains the favored handedness of the helix. That is, a right-handed helix with a twist of +26.7° does not provide favorable contacts between the methyl groups and adjacent *R*-4MeCyCo6 tails (SI, Figure S3). Finally, we note that this model for TAP−*R*-4MeCyCo6 hexad stacks provides an excellent fit between the calculated and experimental diffraction patterns without the introduction of nonzero slide or roll values between adjacent hexads, two additional helical parameters that are shown in Scheme 1.

Fiber Packing Arrangement of TAP−*R*-4MeCyCo6 Hexad Stacks. The equatorial arcs of the TAP−*R*-4MeCyCo6 X-ray fiber diffraction pattern (Figure 3A, *v*) match precisely the series of intensities of hexad stacks organized in a hexagonal lattice that is orthogonal to the fiber axis with an interstack spacing of 33.5 Å (Figure 4). Our fiber diffraction model of TAP−*R*-4MeCyCo6 assemblies is shown in Figure 4 in a hexagonal lattice with 33.5 Å spacing in axial and side-view orientations. The 4-methylhexanoic acid tails of the *R*-4MeCyCo6 units are fully extended in this model. Nevertheless, there is still ca. 8 Å separating the closest approach of carboxylate groups of neighboring stacks (Figure 4C), sufficient distance for the electrostatic charges of neighboring stacks to be screened by intervening cations.

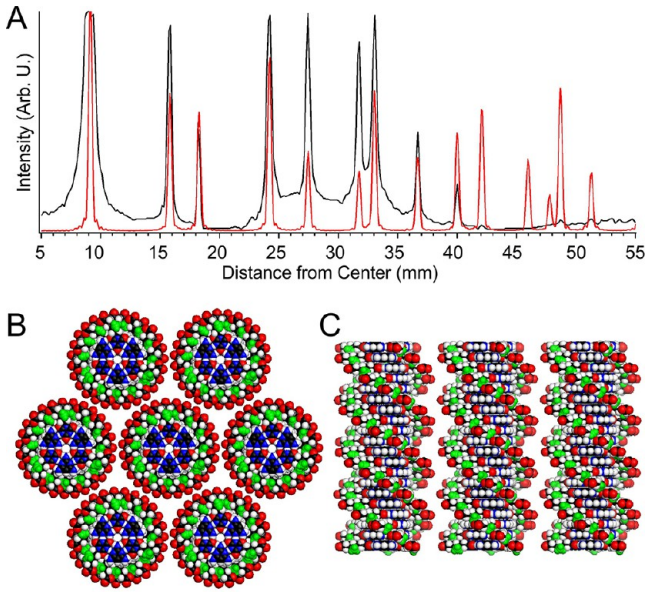


Figure 4. Lattice packing of TAP−*R*-4MeCyCo6 assemblies within fibers. (A) Black trace: Plot of diffraction intensities along the equatorial line of the right-hand side of the pattern shown in Figure 3A. Red trace: Plot of predicted intensities for a hexagonal lattice with a single scatterer at each lattice point and a lattice spacing of 33.5 Å. The excellent correspondence of peak positions confirms the hexagonal packing lattice of TAP−*R*-4MeCyCo6 assemblies. (B) Axial view of hexad stack model structures placed in a hexagonal lattice with a lattice spacing of 33.5 Å. (C) Edge view of three middle hexad stacks of B. Methyl groups of *R*-4MeCyCo6 are shown in green.

Domain Dimensions. The radial widths of equatorial arcs of the TAP−*R*-4MeCyCo6 fiber diffraction pattern provide information about the size of crystalline domains within the

https://doi.org/10.1021/jacs.0c12010
J. Am. Chem. Soc. 2021, 143, 6079−6094

macroscopic fiber by application of the Scherrer equation.[54] In particular, the arcs located at reciprocal distances of 14.5, 7.3, and 6.7 Å were analyzed because their radial intensity profiles did not show signs of broadening due to detector saturation (Figures 4A, 5A). All three arcs were determined to have an average full width at half-maximum of 0.35 mm. This measurement reveals an average crystalline domain size of 700 Å. Considering the interstack spacing within the hexagonal lattice, this average domain size indicates about 400 hexad stacks per crystalline domain.

Stack Tilt Angles within Macroscopic Fibers. The equatorial arcs also provide information on the organization of the crystalline domains within the macroscopic fiber: information that was critical for obtaining a good match between calculated and experimental patterns. Intensity arcs in fiber diffraction patterns are indicative of axial disorder[55,56] and commonly interpreted as disorder of individual fibers, or bundles, in the present case (Figure 5B), about the main fiber axis. In many cases such intensity arcs can be reproduced by calculating the pattern of a single fiber and then adding to this pattern rotated versions of the same pattern that are intensity-weighted with a Gaussian function that is centered about the vertical axis. This approach was initially explored but did not produce patterns that satisfactorily matched the experimental pattern. We then considered that the most populated angle of crystalline bundles might not be collinear with the main fiber axis. To test this hypothesis, a series of diffraction patterns was generated in which the axis of a single stack was tilted relative to, and precessed about, the *z*-axis (the axis vertical to the image plane and perpendicular to the X-ray beam) (Figure 5B). The pattern produced for each stack tilt angle was itself the summation of a series of patterns calculated for an evenly spaced set of rotational angles about the *z*-axis. Patterns generated by this approach with tilt angles between 5° and 10° are much more similar to the experimental pattern than that of the perfectly orientated stack (i.e., 0° tilt) or that with simulated Gaussian-weighted disorder. For example, the pattern calculated for the 6° tilt is shown in Figure 5D.

To obtain a quantitative estimate for the distribution of stack-bundle tilt angles, the intensity profiles of the arcs at reciprocal distances of 14.5, 7.3, and 6.7 Å were fit by a function that was the weighted sum of the intensity profile of an arc of intensity at a reciprocal distance of 8.8 Å for the series of patterns for tilt angles ranging from 0° to 24°. The shape of the weighting function was constructed to represent a smooth potential (or probability distribution) for which the angular width of the function and most populated angle were varied to obtain the best fit of the calculated arc intensity with the experimental pattern arc intensities. The best-fit curve, shown in Figure 5E, indicates a maximum probable tilt angle at 6° and a full-width at half-maximum variance of tilt angles of around 10° (i.e., tilt angles of 2° to 12° are most populated). The full set of calculated patterns for single stacks of 100 hexads, with tilt angles of 0° to 24°, are provided in the SI (Figure S4). The calculated diffraction pattern that is the weighted sum of these patterns, with the values shown in Figure 5F, is the pattern shown above as Figure 3B.

Fiber Structure of TAP−CyCo6 Hexad Supramolecular Assemblies. The structure of TAP−CyCo6 hexad assemblies was also investigated by X-ray fiber diffraction. The findings are similar to those for TAP−*R*-4MeCyCo6, but the differences are very informative.

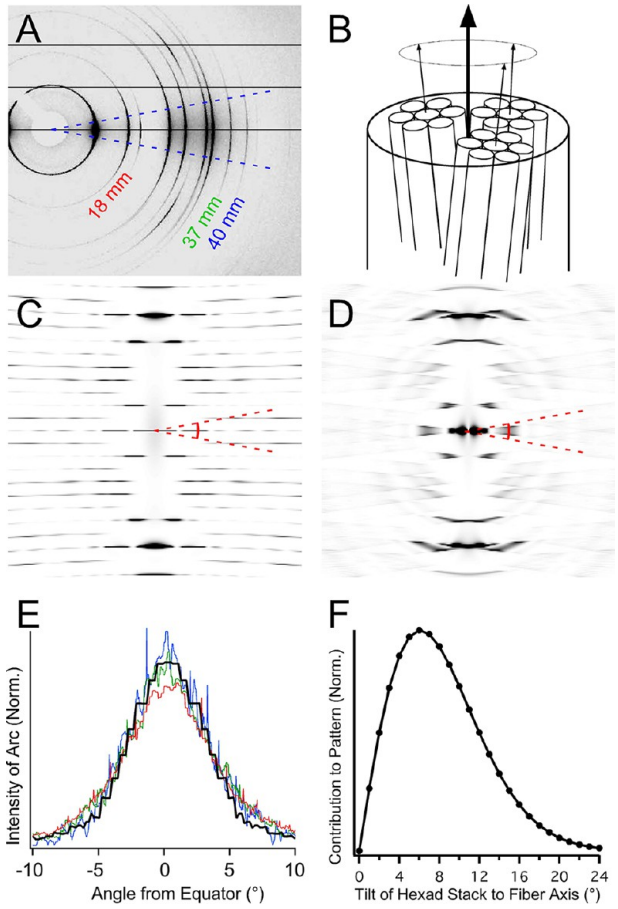


Figure 5. (A) Close-up image of equatorial arcs of the TAP−*R*-4MeCyCo6 pattern used to calculate the angular distribution of hexagonal bundle orientation with respect to the main fiber axis. Specific arcs used for fitting the angular distribution are labeled with their reciprocal distances. (B) Schematic representation of the model for the hierarchy of order within the macroscopic fiber. The angle of individual bundles with respect to the main fiber axis illustrates the origin of axial disorder in the X-ray diffraction pattern. Bundles of stacks are represented as seven hexagonally packed hexad stacks, but the bundle size calculated from arc widths is ca. 400 hexad stacks. (C) Calculated diffraction pattern for the model of TAP−*R*-4MeCyCo6 shown in Figure 3D, with 100 hexads and with a 0° tilt with respect to the vertical axis. (D) Calculated diffraction pattern for the model of TAP−*R*-4MeCyCo6 shown in Figure 3D, with 100 hexads and with a 6° tilt with respect to the vertical axis. Red solid arcs overlaid on calculated patterns C and D show where intensities were measured for determining the weighting function for summation of calculated patterns to best fit the experimental pattern. Patterns shown in C and D are the sums of patterns calculated with the structure rotated, in 10° increments, about the *z*-axis. (E) Thin colored data curves are the intensities of the arcs indicated in panel A, with color coding indicating correspondence with inverse distance labels in panel A. Solid black curve is the RMSD best-fit curve resulting from the summation of arc intensities extracted from calculated patterns for tilt angles of axial rotation ranging from 0° to 24° with the weighting function $f(\theta) = A \cdot \sin(\theta) \cdot \exp(-k\theta^p)$, the shape of which was adjusted using parameters A, k, and p to find the best fit as a function of θ (hexad stack tilt angle). (F) Weighting function for the angle of bundles with respect to the main fiber axis based on the fitting of calculated arc intensities with experimental arc intensities, as shown in panel E.

Hexad Spacing and Helical Twist. The dominant feature of the TAP−CyCo6 fiber diffraction patterns is again a strong

https://doi.org/10.1021/jacs.0c12010
J. Am. Chem. Soc. 2021, 143, 6079−6094

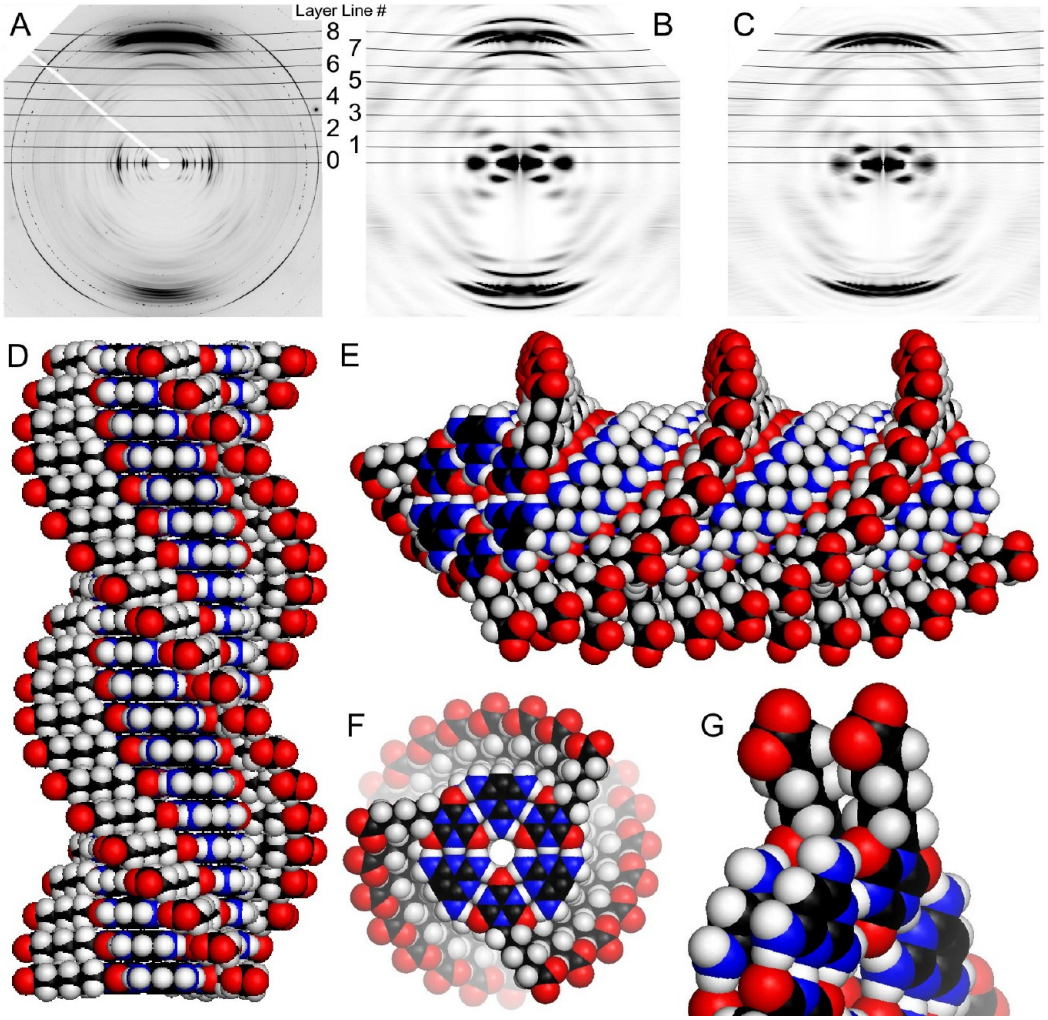


Figure 6. (A) X-ray diffraction pattern of a fiber composed of TAP and CyCo6 in a 1:1 ratio. The orientation of the fiber is vertical relative to the diffraction pattern. The strong reflection located at layer line 8 is at a reciprocal distance of 3.4 Å from the equator and indicative of hexads stacked along the fiber axis. (B) Calculated diffraction pattern based on a 100-hexad stack with a helical twist of −15° per hexad and with the hexanoic acid tails fully extended in the plane of the hexads. The pattern shown is the summation of patterns generated by the summing of individual patterns generated from a 100-hexad stack that was precessed about the *z*-axis with variable angles of tilt from the *z*-axis. Additional details on the generation of this pattern, including weighting of patterns with different stack tilt angles, are provided in the text. (C) Pattern generated the same way as the pattern shown in B, except using a model hexad stack in which the hexanoic acid tails were rotated 45° out of the plane of their associated hexads by rotation about the C6−Cζ bond that connects each cyanuric acid with a hexanoic acid tail. (D−F) Three views of an idealized model of TAP−CyCo6 assemblies with the helical twist of −15° per hexad. (G) Close-up image of the hexanoic acid of two adjacent hexads from the model structure shown in D−F.

near-meridional intensity at a reciprocal distance from the equator of 3.4 Å, likewise indicative of hexads stacked along the fiber axis (Figure 6A). However, in this case, the vertical spacings of intensity arcs are best fit by eight layer lines, with the arc at reciprocal distance of 3.4 Å residing on the eighth layer line (Figure 6A). With each TAP−CyCo6 hexad having 3-fold symmetry the fitting of this pattern by eight layer lines can indicate eight hexads per 120° of helical rotation (a helical twist of 15° per hexad), eight hexads per 240° of helical rotation (a helical twist of 30° per hexad), or eight hexads per 360° of helical rotation (a helical twist of 45° per hexad). Thus, analysis of additional features of the diffraction pattern was again necessary to determine which of the possible twist angles corresponds to the actual helical twist of the assemblies. Briefly, the second most intense arc of the TAP−CyCo6 fiber diffraction pattern lies on the seventh layer line, which is the position predicted by the simulated pattern of a model TAP−CyCo6 stack with a helical twist of 15° per hexad (Figure 6B). In contrast, this arc is located on a different layer line for patterns generated for models with helical twists of 30° or 45° per hexad (SI, Figure S5). Additionally, the lighter arcs on the sixth layer line of the calculated TAP−CyCo6 fiber diffraction pattern (Figure 6B) are also matched by the calculated pattern for the model stack with a helical twist angle of 15°, but not by the patterns generated for models with helical twists of 30° or 45° (SI, Figure S5).

In Figure 6, we present three views of a model for the helical structure of TAP−CyCo6 supramolecular assemblies based on a rise of 3.4 Å per hexad and a helical twist of −15° per hexad. The close packing of the CyCo6 hexanoic tails of adjacent hexads provides a rationale for why this twist angle is favored. In contrast, an angle of −15° would be unfavorable for the TAP−*R*-4MeCyCo6 hexad stacks discussed above, as there would be a steric clash between the methyl groups and

https://doi.org/10.1021/jacs.0c12010
J. Am. Chem. Soc. 2021, 143, 6079−6094

neighboring hexanoic tails. Because both TAP and CyCo6 are achiral, a stack of TAP−CyCo6 hexads with a helical twist of +15° is equivalent to a stack with a helical twist of −15°. Therefore, it is not possible to determine the handedness of the helices with TAP−CyCo6 fibers, and both handednesses likely coexist as macroscopic homochiral domains, as we have shown by CD imaging for assemblies in the gel state.[41]

Fiber Packing Arrangement of TAP−CyCo6 Hexad Stacks. The spacing of arcs along the equatorial line of the TAP−CyCo6 diffraction pattern also indicates crystalline order perpendicular to the fiber axis, but in this case packing is in a centered rectangular lattice with a = 37.9 Å and b = 27.7 Å (Figure 7A). This lattice corresponds to two close-approach interstack spacings of 27.7 and 23.5 Å (Figure 7B). Considering that a space-filling model of the TAP−CyCo6 hexad stack is a triple helix with a maximum diameter of about 32 Å (Figure 6F), it was surprising to find these assemblies packed in a lattice with interhelix spacings several angstroms shorter than this diameter. Nevertheless, there is at least one packing configuration that is compatible with the packing of idealized TAP−CyCo6 hexad stacks with a helical twist of 15° in this specific lattice, albeit with some local adjustments in tail geometries necessary to alleviate steric clash (see next section).

Ridges-in-Grooves Packing Stabilizes a Rectangular Lattice. Figure 7B shows a model of seven hexad stacks packed in the centered rectangular lattice deduced from the fiber diffraction pattern. An axial view suggests extensive overlap of neighboring hexad stacks, for both the closer interstack spacing of 23.5 Å and the further interstack spacing of 27.7 Å. However, a side-view of three helices spaced with the lattice distance of 27.7 Å shows that the ridges of a stack (generated by the extended tails) can nest within the grooves (the space between the ridges) of its neighboring stacks (Figure 7C and E). This ridges-in-grooves packing motif is similar to that observed for the packing of α-helices in proteins.[57] In an idealized model, with fully extended CyCo6 hexanoic acid tails, the close approach of the carboxylate oxygen atoms have a center-to-center spacing of 2.8 Å. Considering that the van der Waals radius of an oxygen atom is 1.52 Å (ionic radius is 1.26 Å) this close approach may not constitute a steric clash, particularly if the alkyl chains of the hexanoic acid tails are not fully extended. Also, it is likely that counterions mediate the electrostatic repulsion between carboxylate groups in the fiber state.

The packing of hexad stacks with the shorter interstack distance of 23.5 Å is more complicated. The 3-fold rotational symmetry of the hexads, the integer number of hexads per helical turn (i.e., 24), and the quasi 6-fold symmetry about each stack within this centered rectangular lattice allow each hexad stack to maintain a ridges-in-grooves packing with its six nearest neighbors (Figure 7A), if all seven helices are of the same handedness, have a helical twist of 15°, and are in register with respect to helical twist angles. The main difference between the helices spaced with an axis-to-axis distance of 23.5 Å and those spaced at 27.7 Å is that the ridges of the more closely spaced stacks penetrate deeper into the grooves of the neighboring stack (Figure 7D and F). In this packing configuration the idealized helical stacks with fully extended hexanoic acid tails show steric clashes between two carboxylate groups of neighboring helices at every point where the ridges of one stack transverse the ridges of a neighboring stack. These clashes are at the level of the complete eclipse of the carboxylate groups (Figure 7G).

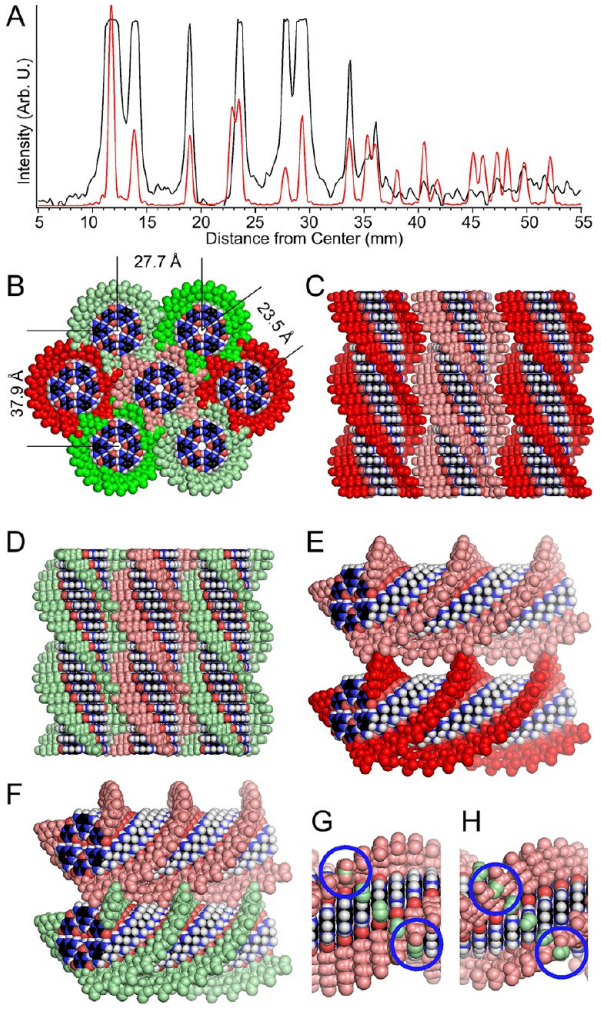


Figure 7. Model for the lattice packing of TAP−CyCo6 assemblies within fibers. (A) Black trace: Plot of diffraction intensities along the equatorial line of the right-hand side of the pattern shown in Figure 6A. Red trace: Plot of predicted intensities for a centered rectangular lattice with single scatter at each lattice point and lattice constants of a = 37.9 Å and b = 27.7 Å. The excellent correspondence of peak positions confirms the packing lattice of TAP−CyCo6 assemblies. (B) Axial view of hexad stack model structures placed in a centered rectangular lattice with lattice constants of a = 37.9 Å and b = 27.7 Å. The close interstack spacing of 23.5 Å is also emphasized. To illustrate the overlap of the circumferences traced by these stacks, the hexanoic tails of neighboring stacks are depicted in different shades of green or red. (C) Edge view and skew view (E) of horizontally orientated hexad stacks of B with a center-to-center spacing of 27.7 Å and the same color scheme. (D) Edge view and (F) skew view of diagonally orientated hexad stacks with center-to-center spacing of 23.5 Å and the same color scheme. (G) View illustrating the most serious steric clashes that appear between the helices shown in F. In this representation the color scheme is the same as in F. Only the carboxylate groups of the lower helix of F with light green ridges are shown. Blue circles emphasize the two groups with most serious steric clash, which is complete overlap with carboxylate groups of the stack with pink ridges. (H) Same view as F except that the C6−Cζ bond between the cyanuric acid group and the hexanoic acid tails is rotated so that the fully extended tails are 45° out of the plane of the hexads. Note that with tails in this configuration the steric clash between carboxylate groups is reduced by allowing the hexanoic acid tails to interdigitate.

https://doi.org/10.1021/jacs.0c12010
J. Am. Chem. Soc. 2021, 143, 6079−6094

As mentioned above, shortening of the hexanoic acid tails by rotations about alkane C−C bonds can reduce this overlap. A change in the angle at which these tails extend from the hexad stack could also eliminate this steric clash. The idealized helical structure shown in Figure 6 and used for the lattice model presented in Figure 7B has the extended hexanoic acid tails lying in the same plane as their respective hexads (Figure 6G). The rotation of these tails out of the plane of the hexads, such as by rotation about the bond between C6 of the cyanuric acid group and the Cζ of the hexanoic acid tails, can place the carboxylate groups of neighboring hexad stacks so that they interdigitate (Figure 7H), which replaces unfavorable steric clash with favorable van der Waals contacts. The theoretical diffraction pattern for a 100-hexad stack with the tails rotated to a 45° angle from the hexad plane was examined to determine if this structure reproduces the experimental diffraction pattern. The calculated pattern, shown in Figure 6C, has most of its distinguishing features in common with the pattern calculated with the tails in the plane of the hexads (Figure 6B). However, this model with tails rotated out of the plane of the hexads results in the loss of the arc of intensity that is seen on the seventh layer line for the model with tails within the hexad planes and, significantly, also in the experimental pattern. Thus, while the local rearrangement of the hexaonic acid tails may be necessary to alleviate steric clashes, the arc observed along the seventh layer line suggests that most of the hexaonic acid tails lie within the plane of their respective hexads.

It is important to note that the lattice packing model shown in Figure 7 is constructed solely from left-handed hexad stacks, but a model constructed from right-handed helices is equally consistent with the observed diffraction pattern. Indeed, domains of hexad stacks of both handedness likely coexist within the macroscopic fibers. Finally, our attempts to create lattice packing models without steric clash that involve a mix of right-handed and left-handed helices were less successful, and we believe that packing models constructed with helices of the same handedness are more realistic based on our previous observation that TAP−CyCo6 assemblies have a very strong propensity to form homochiral macroscopic domains when solutions of these assemblies form hydrogels. We expect the fiber state to be structurally similar to that of the homochiral domains of the gel state, which would be the state of the mother liquor just prior to the formation of macroscopic fibers.

QM Studies Reveal a Shallow Twist Angle Potential for Stacked Hexads. QM methods were used to explore the intrinsic intermolecular interactions contributing to the structure and stability of the hexad assemblies using two model systems, TAP with CyCo6 and TAP with Cy (unmodified cyanuric acid). Here, we focus on results from computational analyses of interaction energies associated with variations in the twist and roll degrees of freedom between stacked hexads (Scheme 1), as these helical parameters are the most pertinent to understanding experimental observations (see below). The results from analyses of the rise and slide parameters are reported in the SI.

Figure 8A shows the QM interaction energy (IE) as a function of twist angle for the TAP−Cy system with two hexads computed using coupled cluster theory through perturbative triples [CCSD(T)], second order Møller−Plesset (MP2) perturbation theory, and symmetry adapted perturbation theory (SAPT). Detailed comparisons between the computational results are available in the SI. Overall, the

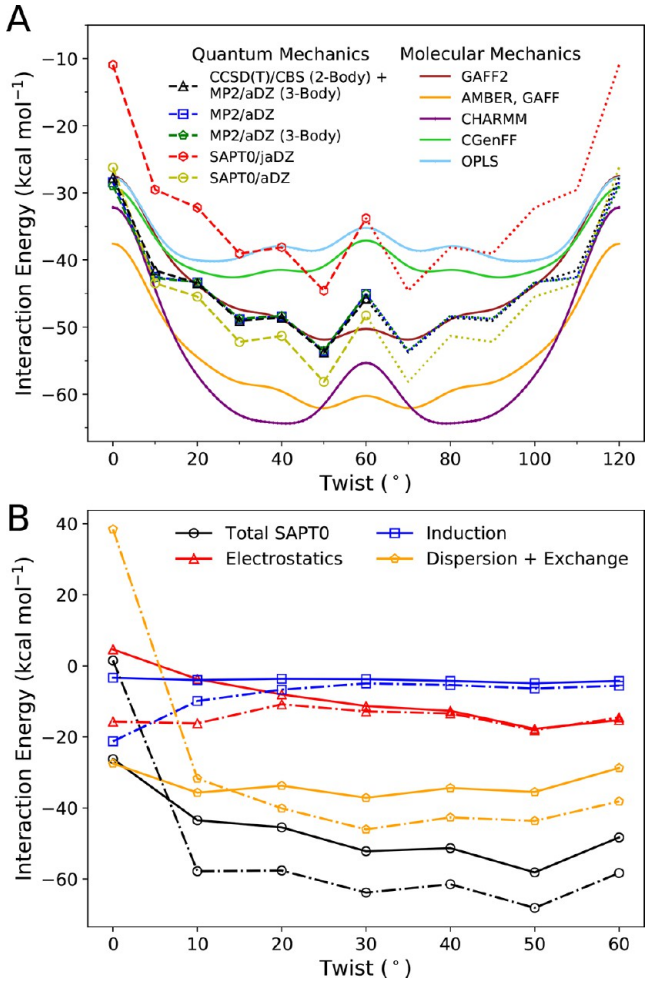


Figure 8. (A) Stacking interaction energy at different twist angles for two hexads of the TAP−Cy system. The rise value between hexads is 3.4 Å. Geometries with twist angles from 120° to 60° are the mirror images of those with twist angles from 0 to 60°. Dotted lines are extrapolations of the data between 0° and 60°. (B) SAPT0/aug-cc-pVDZ interaction energies and their components for the TAP−CyCo6 system (dashed) and the TAP−Cy system (solid). At 0°, the exchange repulsion caused by extensive overlap of CyCo6 tails results in high-energy configuration.

potential energy profiles across the various QM methods show a greater than 20 kcal mol^{-1} increase in stability arising from twisting the hexads away from an initial twist angle of 0°, which corresponds to like heterocycles in register (i.e., perfectly eclipsed), with modest IE fluctuations and a preference for larger twist angles.

We explored the physical nature of the intermolecular interactions using the many-body expansion (MBE) and SAPT approaches. The MBE method separates the IE into pairwise and higher-order nonadditive contributions. Results at the MP2 level (Figure 8A and Tables S1−S7) show that pairwise interaction between the heterocycles in the two stacked hexads accounts for most of the total IEs, which range from −28 to −54 kcal mol^{-1}. Contributions involving heterocycle triples are slightly repulsive and account for 2 to 4 kcal mol^{-1}. Contributions beyond three-body effects are small and account for less than 0.5 kcal mol^{-1}.

The SAPT approach partitions the IE into physically meaningful components (electrostatics, exchange-repulsion, induction/polarization, and London dispersion). Figure 8B

https://doi.org/10.1021/jacs.0c12010
J. Am. Chem. Soc. 2021, 143, 6079−6094

shows the SAPT0 twist angle IE profile and analyses for both the TAP−Cy and the neutral (carboxylic acid form) TAP−CyCo6 systems (CyCo6 anions were used for the MD simulations presented below). Except for the IE at 0° twist angle, the total IEs are qualitatively similar in the absence or the presence of the carboxylic acid tail. This result implies that formation of a twisted, helical structure is driven, in part, by favorable electronic interaction of the heterocycles of the hexad core, in addition to other contributions by the exocyclic tails. A more favorable energy for twist angles away from a perfect eclipse of rings is similarly observed for base-steps within duplex DNA, which show significant improvement of the IE upon twisting even in the absence of a backbone.[58]

For most twist angles, the SAPT computations show that the induction/polarization contribution to the IE is rather small, whereas the van der Waals terms (attractive London dispersion terms plus repulsive exchange terms) are the most significant in determining the twist angle potential, with increasingly favorable dispersion interaction between the aliphatic carbon chains at angles of 20° or greater. For lower twist angles, where the tails approach each other closely, van der Waals interaction is less favorable due to higher steric repulsion, while electrostatic interaction is favorable (for hexads with protonated, charge-neutral CyCo6 hexanoic acid tails). The unfavorable tail−tail interaction quickly diminishes beyond a 10° twist, and therefore low twist angles are energetically accessible. Furthermore, conformational flexibility of the tails can alleviate tail−tail repulsion. This was not explored in the QM calculations for computational practicality.

The SAPT computations show a rapid increase in energy if one hexad is tilted with respect to another (the roll coordinate, Figure S6). This energy increase is consistent with the observation of long fibers with large persistence length,[22,26] as discussed below in the context of MD simulations and AFM fiber images. The rise and slide degrees of freedom have also been explored using SAPT computations and are reported in the SI (Figure S7 and Table S8).

Classical MD Simulations Reproduce the Stability of Isolated Stacked Hexad Supramolecular Polymers in Water. MD simulations were used to investigate aspects of the structure of TAP−CyCo6 assemblies in solution that are not accessible by fiber diffraction analysis or QM computations. We used the AMBER, GAFF, and GAFF2 force fields, which showed the best agreement with quantum mechanical results on hexad stacking interaction energy (see Figure 8A and additional details in the SI). The simulations were performed on a single stack of 20 hexads. Figure 9A shows a snapshot of the TAP−CyCo6 assemblies extracted from the MD simulations.

Figure 9B shows the distributions of the twist and roll angles (Scheme 1) over the MD trajectories for the three different force fields. The twist angle distribution shows strong preference for twisting the hexads away from an eclipsed state (0°), which is consistent with the significantly more favorable QM IEs shown in Figure 8. The three force fields have varying twist angle distributions, but all agree with the QM trends of preference for high twist angles (30−60°) over low twist angles (0−30°). Structural analyses for the other degrees of freedom and detailed structure distributions for each hexad in the stacks are available in the SI (Figures S8−S11).

All force fields show a preference for maintaining the hexad at a parallel configuration, as tilting the plane of one hexad with respect to the plane of an adjacent hexad along either perpendicular axis in the plane of the hexads leads to a rapid increase in steric repulsion for even modest nonzero roll angles, which is in agreement with the SAPT0 computations of the roll potential energy profile (SI, Figure S6).

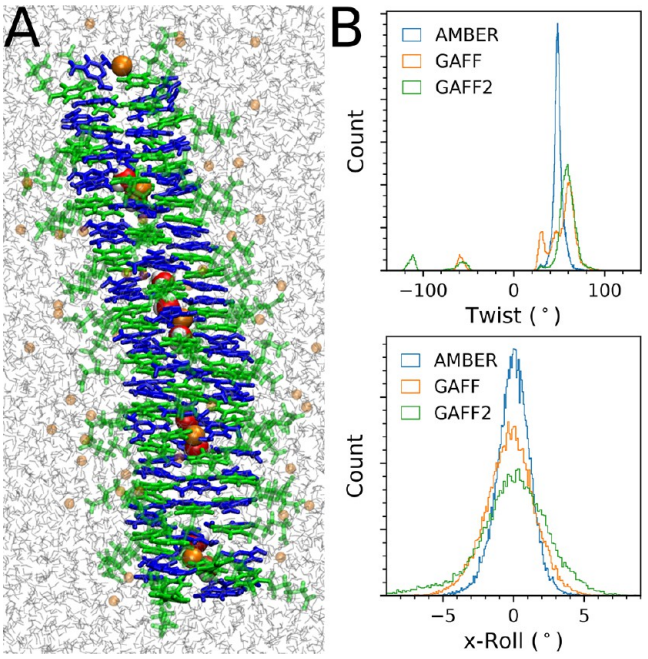


Figure 9. (A) Representative snapshot of the MD simulations. CyCo6 and TAP are depicted in green and blue, respectively. The simulations show that sodium ions and water molecules can be captured within the hexad stacks, as displayed in the space-filling representation. (B) Histograms of the twist and roll angles extracted from MD simulations utilizing the AMBER, GAFF, and GAFF2 force fields. Histograms of the remaining structural parameters are available in the SI.

MD Simulations Show Trapping of Water and Sodium Ions within Hexad Stacks. MD simulations of TAP−CyCo6 supramolecular polymers exhibit trapping of sodium ions and water molecules even though the starting configurations do not include them within the assembly (Figures 9A and S12). During the 10 ns simulation, sodium ions show limited mobility between hexads. Interaction between solvent ions and molecules within hexameric rosettes has been proposed before. MD simulations showed that chloroform molecules can occupy the central pore in the hexads in chloroform solution.[10] Experimental evidence for the capture of water molecules within stacked hexads was reported for bicyclic self-pairing molecules that create an axial channel that is considerably larger than for the stacked TAP−CyCo6 hexads.[59,60] More recently, Petelski and Guerra used density functional theory to explore the interaction of various cations and anions with both single and doubly stacked melamine hexameric rosettes.[61] Asanuma and co-workers considered that hexad-based oligomers could function as ion channels.[42] Our results are consistent with this suggestion and highlight the similarity between hexad supramolecular polymers and stacked guanosine-containing polymers, which incorporate cations in their central pores.[62]

Computational Studies of Hexad Fiber Persistence Length. Beginning with the first report of water-soluble TAPAS and Cy supramolecular polymers,[22] it has been

https://doi.org/10.1021/jacs.0c12010
J. Am. Chem. Soc. 2021, 143, 6079−6094

apparent that individual fibers have an extremely long persistence length (Figure 10A,B). Qualitatively, persistence length is a measure of polymer stiffness, being defined as the contour length over which the directions of two vectors tangent to the polymer axis become uncorrelated with each other.[63] Results from the MD simulations are consistent with the long persistence length of hexad polymers and confirm the strong penalty associated with the introduction of a nonzero roll angle between adjacent hexads within a stack. Analysis of 3×10^4 roll values from the AMBER force field trajectory indicates a most probable roll angle of slightly less than 1° and a mean roll value of 1.4° (SI, Figure S13). This distribution of roll angles and the interhexad spacing of 0.34 nm predict a persistence length for hexad stacks of approximately 800 nm (SI). The same calculation based on MD simulations with the GAFF and GAFF2 force fields yields predicted persistence lengths of approximately 400 and 250 nm, respectively. For comparison, the persistence length of duplex DNA is approximately 50 nm.[64] Consistent with the much greater calculated persistence length of TAP−CyCo6 fibers, DNA in the AFM image presented in Figure 10 shows much greater curvature than hexad stacks over the same polymer contour length.

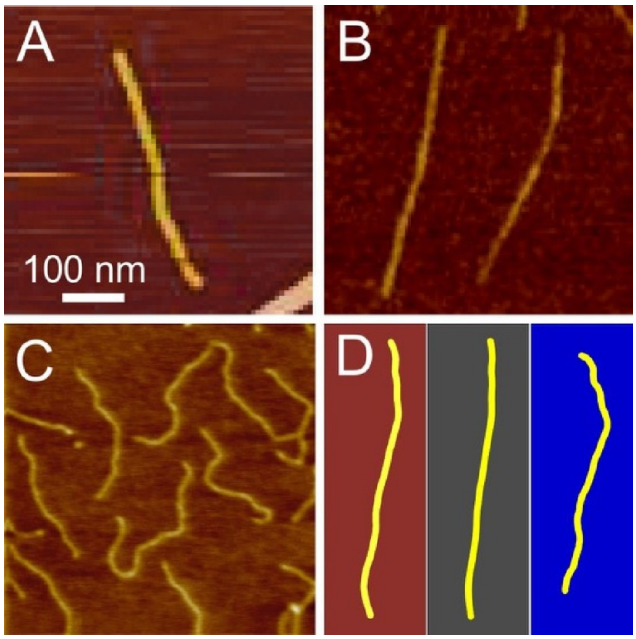


Figure 10. Analysis of the persistence length of TAP−Cy assemblies. (A) Individual TAPAS−Cy and (B) TAP−CyCo6 fibers obtained with AFM. These individual fibers are extracted from the experimental data shown in Figure 2. (C) Individual fibers of double-stranded DNA (reprinted from ref 66; copyright 2014, with permission from Elsevier). (D) Simulated fibers generated using the AMBER roll angle distributions scaled by (brown) 1, (gray) 0.5, and (blue) 2. All panels have the same 100 nm scale. Additional simulated fibers are provided in the SI (Figure S15).

We also investigated how well hexad roll angles measured in MD simulations correlate with experimentally observed polymer persistence lengths by comparing the contour paths of model polymers generated from MD-derived roll angles with the contour paths of actual hexad stacks observed in AFM images. Briefly, hexad stacks were modeled as a variation of Gaussian polymers[65] in which the energy potential for changing the axial direction of a polymer depends only on the energy associated with changes in roll angles between adjacent hexads. An example of a resulting model polymer contour path based on the distribution of AMBER force field roll angles is shown in the left panel of Figure 10D. For comparison, the middle and right panels of Figure 10D show results for simulation with AMBER roll angle distributions scaled by 0.5 and 2, respectively. The predicted curvatures in the three panels deviate minimally from those shown by the AFM images in Figure 10A and B, revealing that the roll angle distribution predicted by the AMBER force field reliably generates assemblies with average curvature similar to that observed experimentally. Thus, the energetic potential, in the solution phase, for the roll angle defined between adjacent hexads of a linear TAP−CyCo6 assembly is accurately reproduced by the AMBER force field.

DISCUSSION

The intimate structural details of supramolecular polymers, particularly those formed in aqueous solution, are both of great interest and very difficult to obtain. The properties of these polymers are, of course, determined by their structure, and their properties are of significant practical and theoretical importance. In this report we presented the results of extensive XRFD and computational analyses of two polymer systems, one chiral and one achiral, that reveal exquisite features of the association, stacking, twisting, and packing of these systems. These polymers are of additional interest because they are possibly relevant to the formation and chemical evolution of informational (genetic) polymers.

Taken together, the X-ray fiber diffraction data and computational results reported here reveal and explain the formation of stiff, long, noncovalent assemblies of twisted stacks of essentially planar hexameric rosettes for polymers formed from TAP and a substituted cyanuric acid. Stacked hexameric rosettes of such molecules have been considered as a possible motif for supramolecular assembly for three decades, and water-soluble linear structures consistent with the formation of such assemblies were reported several years ago, but direct evidence in support of the stacked hexad model has been absent. The results reported here are the first direct, extensive confirmation of this model. Analyses of the TAP−*R*-4MeCyCo6 and the TAP−CyCo6 assemblies reveal helical hexad stacks with well-defined but different twist angles and show that the twist angle varies with exocyclic modifications and the local environment.

The QM calculation results reveal that the stacking interaction of the hexads with twist angles between 30° and 60° surpasses that of most DNA base-pair steps. The DNA bases have a larger average surface area (due to the purine nucleobases), but the higher symmetry of the hexads allows all the bases to remain in proximity upon twisting. Recent values for DNA stacking energies range from −11.7 to −18.4 kcal mol^{-1},[67] which is about 30% of the stacking energies for these hexads. The greater number of bases in the hexads allows more pairwise interactions between the bases compared to the DNA base-pair steps (although some of them are longer-range and therefore weaker).

The fiber diffraction data show an interplanar distance between stacked hexameric rosettes of 3.4 Å and a flexible twist angle that is apparently controlled by close atomic contacts. TAP−*R*-4MeCyCo6 and the TAP−CyCo6 assemblies adopt different twist angles (−26.7° vs ±15°) because of the

https://doi.org/10.1021/jacs.0c12010
J. Am. Chem. Soc. 2021, 143, 6079−6094

presence of the methyl group on *R*-4MeCyCo6 that is absent from CyCo6. Neither twist angle is the favored result of the QM and MD simulations, which predict high and variable twist angles (30−60°) and a relatively shallow twisting potential except close to the eclipsed angle (0°). The contrast between the fiber diffraction data and the computational results reveals the importance of interstack contacts in governing twist angles in the fiber state. The very close packing of the hexad stacks within the TAP−CyCo6 fibers necessitates the adoption of twist angles that allow for favorable interstack contacts and minimization of steric clashes. The modulation of the twist angle of supramolecular stacked structures by *inter*stack and *intra*stack correlations in the fiber state was previously considered. Peterca et al.[68] examined the stacking geometries of C_3-supramolecular dendrimers, which have 3-fold rotational symmetry analogous to that of the TAP−Cy hexameric rosettes. They reported that in the absence of interstack correlation the twist angle is ca. 60° and that interstack interaction in the fiber state lowers the twist angle to ca. 15° to favor the close packing of assemblies, consistent with the ridges-in-grooves packing model reported here for TAP−CyCo6 assemblies.

The impact of close-packing on twist angle is also consistent with the observation that TAP−CyCo6 assemblies form macroscopic domains of the same helical twist and exhibit an extraordinary sensitivity to the presence of chiral dopants that induce the formation of exclusively homochiral supramolecular polymers,[41] an observation that suggests transfer of helicity in fibers from one stack to another. Also in support of lattice contacts influencing the value and uniformity of the TAP−CyCo6 hexad twist angle is the observation that circular dichroism (CD) signals are not detected for the TAP−CyCo6 assemblies until the concentration is sufficiently high that the assemblies approach the gel state.[41] This observation is supportive of the findings from QM and MD simulations, which indicate that twist angles are more varied and dynamic for isolated stacks in solution. The more distant spacing of the TAP−*R*-4MeCyCo6 assemblies suggests that interstack interactions may be less critical determinants of the helical twist angle in this case. Nevertheless, the interaction of the methyl groups of hexanoic acid tails of neighboring hexads provides sufficient energetic contributions for TAP−*R*-4MeCyCo6 assemblies to adopt a uniform twist angle in the fiber state.

TAP−*R*-4MeCyCo6 assemblies pack in a hexagonal lattice, whereas TAP−CyCo6 packs in a centered rectangular lattice. A centered rectangular lattice differs from a hexagonal lattice in that any given assembly has four neighboring assemblies at a close spacing (e.g., 23.5 Å) and two neighboring assemblies at a greater spacing (e.g., 27.7 Å) (Figure 7B), whereas all six neighboring assemblies are at equal spacings in a hexagonal array. The coexistence of two interassembly spacings in the centered rectangular lattice is of potential significance for the packing of TAP−CyCo6 stacks, which can adopt left-handed or right-handed helical twists. Peterca et al. recognized that a lattice with hexagonal symmetry cannot accommodate the close packing of stacks with mixed handedness (i.e., right- and left-handed helices), but that a lattice with rectangular symmetry (i.e., 2-fold or 4-fold symmetry) could accommodate the close packing of helices with different handedness. Accordingly, we attempted to construct clash-free lattices of TAP−CyCo6 stacks with alternating +15° and −15° helical twists within the centered rectangle. The steric clashes were somewhat more severe than those observed for the uniform handedness model, and we did not find any phasing arrangements of stacks in this lattice of mixed handedness that were free of steric clash between side chains.

Both the experimental and computational results clearly show that supramolecular polymers of stacked TAP−CyCo6 hexads have persistence lengths at least an order of magnitude greater than that of duplex DNA. The QM computations show that introducing a nonzero roll angle between stacked hexad rosettes results in sizable steric repulsion. Clearly, the greater symmetry and larger surface area of the hexad units compared to canonical DNA base-pairs render any small bending sterically and energetically unfavorable.

We note that the comparative resistance to roll angle distortion of the stacked hexads may provide an important advantage for their development as the recognition units of synthetic informational polymers and, historically, as a scaffold for the pre-enzyme formation of early genetic polymers. It is frequently observed that cyclization competes with chain extension in the nonenzymatic polymerization of RNA and DNA nucleotides.[69] The stiffness of stacked TAP−CyCo6 hexads, and similar hexad-based assemblies, will inhibit cyclization and facilitate their efficient spontaneous conversion to long covalent polymers.

CONCLUSIONS

We report results from examining the structure and the noncovalent interactions of the supramolecular polymers formed by $(TAP{-}Cy)_3$ hexameric rosettes using two powerful, complementary techniques: X-ray fiber diffraction analysis and QM and MD simulations. We previously reported that aqueous solutions of equimolar mixtures of TAP and Cy monomers, above the minimal assembly concentration, self-assemble into gel-forming supramolecular polymers. XRFD analysis of macroscopic fibers of $(TAP/CyCo6)_3$ or $(TAP/R\text{-}4MeCyCo6)_3$ polymers provides direct experimental evidence for stacked hexameric rosette structures of supramolecular polymer assemblies. The data reveal highly ordered, rigid, twisted assemblies within a tightly packed crystal structure. The presence of the methyl substituent in 4MeCyCo6 is shown to favor the formation of assemblies having greater twist angles compared to assemblies containing CyCo6. Comparisons between the results of the X-ray diffraction analysis of hexad stacks in the fiber state and computational analysis of an isolated hexad stack in the solution state highlight the effects of inter- and intrastack correlation in modulating the helical structure of the assemblies and in determining the favored twist angle. Noncovalent interactions (particularly dispersion) and solvation effects (exemplified by the capture of cations and water molecules within the stacks) are shown to contribute to the stability of the assemblies. MD simulations show that the stacked rosettes have a small mean roll angle, which accounts for the extremely long persistence lengths observed in AFM images of these fibers. These findings provide a rationale for the remarkable properties observed for the TAP−Cy noncovalent hexameric rosette assemblies.

METHODS

Quantum Chemical Calculations. Geometries of single hexameric rosettes composed of alternating TAP and Cy bases or TAP and CyCo6 bases were optimized, with the Q-Chem package (version 4.4),[70] using the B97-D3 dispersion-corrected density functional approximation[71,72] and the cc-pVDZ basis set.[73]

https://doi.org/10.1021/jacs.0c12010
J. Am. Chem. Soc. 2021, 143, 6079−6094

Counterpoise-corrected,[74] density-fitted, second-order Møller–Plesset perturbation theory[75] with the aug-cc-pVDZ basis set[76] was used to compute the total IE between the two optimized hexads. In addition to the MP2 IEs computed between the complete two hexads, we also explored the many-body expansion as an alternative, more scalable approach (more details in the SI).

A more accurate two-body approximation to the IE was also computed using a focal point approach.[77−80] In this approach, the energy is computed with coupled-cluster theory through perturbative triples and extrapolated to the complete basis set limit using Helgaker's two-point scheme[81] with the aug-cc-pVTZ and aug-cc-pVQZ basis sets. The coupled-cluster energy is computed using the density-fitted, frozen natural orbital approximation.[82]

To analyze the noncovalent interactions in the system, IEs between two hexads of TAP−CyCo6 with neutral CyCo6 units and TAP−Cy hexads at various configurations were computed using symmetry-adapted perturbation theory[83] with the jun-cc-pVDZ and the aug-cc-pVDZ basis sets.[84] All quantum chemical IEs were computed using the Psi4 package (version 1.2).[85] The SAPT0/aug-cc-pVDZ computations were performed with a development version of Psi4.

The QM twist-angle potential energy profiles were compared with profiles computed using various classical force fields (see below).

Molecular Mechanics Calculations. IE and MD simulations were conducted using various widely used force fields. When nonstandard parameters were required, a variety of tools were employed to generate them. Merged MMFF and CHARMM force field parameters (labeled "CHARMM" in Figure 8A) were obtained using the SwissParam Web server.[86] Parameters for the CHARMM General Force Field (CGenFF) were obtained using the CGenFF Web server.[87] OPLS-AA parameters were obtained using the LibParGen Web server.[88] Parameters for AMBER, the General AMBER Force Field (GAFF),[89] and the second-generation GAFF, or GAFF2 (version 2.1), were obtained using the Antechamber module of the AmberTools package.[90]

Molecular Dynamics Simulations. The OpenMM software library (version 7.2)[91] was employed to perform the simulations. The starting stacked structures were constructed from the QM-optimized geometries for one hexad. Twenty hexads were initially separated by 3.4 Å, and the twist angle between each adjacent pair of hexads was set to 30°, in simulation boxes of approximately 60 Å along the x- and y-axes. Along the z-axis (the polymer axis), the box was extended an additional 30 Å beyond the terminal hexads. Enough sodium ions were added to neutralize the system (i.e., one per CyCo6), and thousands of TIP3P water molecules[92] were added to fill the box. Nonbonded interactions were cut off at 15 Å, and a switching function was applied at 13 Å. Particle Mesh-Ewald summation[93] was applied for the electrostatic terms. The Langevin integrator[94] with 1 fs time step and 1 ps^{-1} friction coefficient was used. After energy minimization, the system was heated gradually to 300 K in 100 ps and then the temperature was maintained at constant volume for another 100 ps, followed by equilibration at a constant pressure of 1 bar for 125 ns. The production simulation in isobaric, isothermal conditions lasted 10 ns with snapshots saved every 5 ps.

Structural Analysis. The geometric configuration of the systems along the degrees of freedom shown in Scheme 1 was determined by computing the translation and rotation matrices that minimize the least root mean squared displacement between the two hexads (full details in the SI).

Persistence Length Analysis. In the persistence length computations, we were concerned with the relative roll between adjacent hexads along any direction, and not around a particular axis. Therefore, to obtain the roll angles from the MD simulations, we fit the QM-optimized hexad to each one of the hexads extracted from the MD trajectories. Then, the roll angles were computed as the angle between the two vectors normal to the plane of the hexad (more details in the SI).

Fiber Diffraction. Fibers were prepared by suspending a 5−10 μL drop of fibril suspension between two glass rods, silanized and sanded at the tips, approximately 1.5 mm apart.[95] The fibers were allowed to dry for several days in a closed chamber under 86% relative humidity in equilibrium with saturated potassium chloride. Diffraction data were collected at beamline 4-2 at the Stanford Synchrotron Radiation Laboratory. Fibers were dusted with calcite, and specimen-to-detector distances were determined from the 012 calcite diffraction ring at 3.8547 Å resolution and the 104 ring at 3.0355 Å.[96] Diffraction patterns were analyzed using the program WCEN[97] to determine experimental parameters and positions of reflections.

Fiber Diffraction Analysis. A computer code for simulating fiber diffraction patterns was written in Python. The program used a first-principles approach to calculate the intensity and phase of X-ray waves from the heavy atoms of hexad stacks at each pixel of a flat 2D detector. Atomic numbers were used as scattering factors. An X-ray wavelength of 0.7749 Å and a distance from sample to detector of 338.4 mm, based on the experimental diffraction parameters, were used for the simulations. The fiber diffraction code is available on GitHub at https://github.com/GT-NucleicAcids/fiber-diffraction.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.0c12010.

Procedures, tables and figures associated with XRFD structure determination, and computational studies (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Nicholas V. Hud − *School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332-0400, United States; NSF-NASA Center for Chemical Evolution, Atlanta, Georgia 30332-0400, United States;* orcid.org/0000-0001-7711-6472; Email: hud@gatech.edu

C. David Sherrill − *School of Chemistry and Biochemistry, Center for Computational Molecular Science and Technology, and School of Computational Science and Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0400, United States; NSF-NASA Center for Chemical Evolution, Atlanta, Georgia 30332-0400, United States;* orcid.org/0000-0002-5570-7666; Email: sherrill@gatech.edu

Authors

Asem Alenaizan − *School of Chemistry and Biochemistry and Center for Computational Molecular Science and Technology, Georgia Institute of Technology, Atlanta, Georgia 30332-0400, United States; NSF-NASA Center for Chemical Evolution, Atlanta, Georgia 30332-0400, United States;* orcid.org/0000-0002-0871-664X

Carlos H. Borca − *School of Chemistry and Biochemistry and Center for Computational Molecular Science and Technology, Georgia Institute of Technology, Atlanta, Georgia 30332-0400, United States;* orcid.org/0000-0003-0683-7613

Suneesh C. Karunakaran − *School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332-0400, United States; NSF-NASA Center for Chemical Evolution, Atlanta, Georgia 30332-0400, United States;* orcid.org/0000-0002-9587-4645

Amy K. Kendall − *Department of Biological Sciences and Center for Structural Biology, Vanderbilt University, Nashville, Tennessee 37235, United States*

Gerald Stubbs − *Department of Biological Sciences and Center for Structural Biology, Vanderbilt University, Nashville, Tennessee 37235, United States*

https://doi.org/10.1021/jacs.0c12010
J. Am. Chem. Soc. 2021, 143, 6079−6094

Gary B. Schuster − *School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332-0400, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.0c12010

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was jointly supported by NSF and the NASA Astrobiology Program, under the NSF Center for Chemical Evolution, CHE-1504217, and through research cyberinfrastructure resources and services provided by the Partnership for an Advanced Computing Environment (PACE) at the Georgia Institute of Technology. We thank Prof. Loren Dean Williams for helpful discussions and Prof. Jesse G. McDaniel for helpful discussions and for providing computing time on graphics processing unit resources.

■ REFERENCES

(1) Seto, C. T.; Whitesides, G. M. Self-assembly based on the cyanuric acid melamine lattice. *J. Am. Chem. Soc.* **1990**, *112*, 6409−6411.

(2) Lehn, J.-M.; Mascal, M.; Decian, A.; Fischer, J. Molecular recognition directed self-assembly of ordered supramolecular strands by cocrystallization of complementary molecular components. *J. Chem. Soc., Chem. Commun.* **1990**, 479−481.

(3) Ranganathan, A.; Pedireddi, V. R.; Rao, C. N. R. Hydrothermal synthesis of organic channel structures: 1:1 hydrogen-bonded adducts of melamine with cyanuric and trithiocyanuric acids. *J. Am. Chem. Soc.* **1999**, *121*, 1752−1753.

(4) Lehn, J. M.; Mascal, M.; Decian, A.; Fischer, J. Molecular ribbons from molecular recognition directed self-assembly of self-complementary molecular-components. *J. Chem. Soc., Perkin Trans. 2* **1992**, 461−467.

(5) Zerkowski, J. A.; Seto, C. T.; Whitesides, G. M. Solid-state structures of rosette and crinkled tape motifs derived from the cyanuric acid melamine lattice. *J. Am. Chem. Soc.* **1992**, *114*, 5473−5475.

(6) Bohanon, T. M.; Denzinger, S.; Fink, R.; Paulus, W.; Ringsdorf, H.; Weck, M. Barbituric-acid 2,4,6-triaminopyrimidine aggregates in water and their competitive interaction with a monolayer of barbituric-acid lipids at the gas-water interface. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 58−60.

(7) Koyano, H.; Bissel, P.; Yoshihara, K.; Ariga, K.; Kunitake, T. Effect of melamine-amphiphile structure on the extent of two-dimensional hydrogen-bonded networks incorporating barbituric acid. *Chem. Eur. J.* **1997**, *3*, 1077−1082.

(8) MarchiArtzner, V.; Jullien, L.; GulikKrzywicki, T.; Lehn, J. M. Molecular recognition induced aggregation and fusion between vesicles containing lipids bearing complementary hydrogen bonding head-groups. *Chem. Commun.* **1997**, 117−118.

(9) Ren, Y. Z.; Zhao, B.; Chai, X. D.; Lu, A. D.; Chen, S. G.; Cao, Y. W.; Yang, W. S.; Lu, R.; Duo, J. Q.; Jiang, Y. S.; Li, T. J. Structural study on the alternating Langmuir-Blodgett films deposited with two complementary barbituric acid and melamine derivatives. *Thin Solid Films* **1997**, *293*, 170−174.

(10) Chin, D. N.; Gordon, D. M.; Whitesides, G. M. Computational simulations of supramolecular hydrogen-bonded aggregates - hubM_3, flexM_3, and adamantane-based hubs in chloroform. *J. Am. Chem. Soc.* **1994**, *116*, 12033−12044.

(11) Hanabusa, K.; Miki, T.; Taguchi, Y.; Koyama, T.; Shirai, H. 2-Component, small-molecule gelling agents. *J. Chem. Soc., Chem. Commun.* **1993**, 1382−1384.

(12) Yagai, S.; Nakajima, T.; Karatsu, T.; Saitow, K.-i.; Kitamura, A. Phototriggered self-assembly of hydrogen-bonded rosette. *J. Am. Chem. Soc.* **2004**, *126*, 11500−11508.

(13) Bielejewska, A. G.; Marjo, C. E.; Prins, L. J.; Timmerman, P.; de Jong, F.; Reinhoudt, D. N. Thermodynamic stabilities of linear and crinkled tapes and cyclic rosettes in melamine-cyanurate assemblies: A model description. *J. Am. Chem. Soc.* **2001**, *123*, 7518−7533.

(14) Huck, W. T. S.; Hulst, R.; Timmerman, P.; vanVeggel, F.; Reinhoudt, D. N. Noncovalent synthesis of nanostructures: Combining coordination chemistry and hydrogen bonding. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 1006−1008.

(15) Yagai, S.; Nakajima, T.; Kishikawa, K.; Kohmoto, S.; Karatsu, T.; Kitamura, A. Hierarchical organization of photoresponsive hydrogen-bonded rosettes. *J. Am. Chem. Soc.* **2005**, *127*, 11134−11139.

(16) Yagai, S.; Usui, M.; Seki, T.; Murayama, H.; Kikkawa, Y.; Uemura, S.; Karatsu, T.; Kitamura, A.; Asano, A.; Seki, S. Supramolecularly Engineered Perylene Bisimide Assemblies Exhibiting Thermal Transition from Columnar to Multilamellar Structures. *J. Am. Chem. Soc.* **2012**, *134*, 7983−7994.

(17) Petelski, A. N.; Peruchena, N. M.; Sosa, G. L. Evolution of the hydrogen-bonding motif in the melamine-cyanuric acid co-crystal: A topological study. *J. Mol. Model.* **2016**, *22*, 202.

(18) Mathias, J. P.; Simanek, E. E.; Seto, C. T.; Whitesides, G. M. Design, preparation, and characterization of hydrogen-bonded supramolecular aggregates based on the cyanuric acid melamine lattice. *Macromol. Symp.* **1994**, *77*, 157−166.

(19) Mathias, J. P.; Simanek, E. E.; Zerkowski, J. A.; Seto, C. T.; Whitesides, G. M. Structural preferences of hydrogen-bonded networks in organic solution — the cyclic CA_3•M_3 rosette. *J. Am. Chem. Soc.* **1994**, *116*, 4316−4325.

(20) Mathias, J. P.; Simanek, E. E.; Whitesides, G. M. Self-assembly through hydrogen-bonding - peripheral crowding - a new strategy for the preparation of stable supramolecular aggregates based on parallel, connected CA_3·M_3 rosettes. *J. Am. Chem. Soc.* **1994**, *116*, 4326−4340.

(21) Rakotondradany, F.; Palmer, A.; Toader, V.; Chen, B. Z.; Whitehead, M. A.; Sleiman, H. F. Hydrogen-bond self-assembly of DNA-analogues into hexameric rosettes. *Chem. Commun.* **2005**, 5441−5443.

(22) Cafferty, B. J.; Gallego, I.; Chen, M. C.; Farley, K. I.; Eritja, R.; Hud, N. V. Efficient self-assembly in water of long noncovalent polymers by nucleobase analogues. *J. Am. Chem. Soc.* **2013**, *135*, 2447−2450.

(23) Karunakaran, S. C.; Cafferty, B. J.; Pelaez-Fernandez, M.; Neselu, K.; Schmidt-Krey, I.; Fernandez-Nieves, A.; Schuster, G. B.; Hud, N. V. Exquisite regulation of supramolecular equilibrium polymers in water: chain stoppers control length, polydispersity and viscoelasticity. *Polym. Chem.* **2018**, *9*, 5268−5277.

(24) Chen, M. C.; Cafferty, B. J.; Mamajanov, I.; Gállego, I.; Khanam, J.; Krishnamurthy, R.; Hud, N. V. Spontaneous prebiotic formation of a β-ribofuranoside that self-assembles with a complementary heterocycle. *J. Am. Chem. Soc.* **2014**, *136*, 5640−5646.

(25) Cafferty, B. J.; Fialho, D. M.; Hud, N. V. In *Prebiotic Chemistry and Chemical Evolution of Nucleic Acids*; Menor-Salván, C., Ed.; Springer International Publishing: Cham, 2018; pp 143−174.

(26) Cafferty, B. J.; Avirah, R. R.; Schuster, G. B.; Hud, N. V. Ultrasensitive pH control of supramolecular polymers and hydrogels: pK_a matching of biomimetic monomers. *Chem. Sci.* **2014**, *5*, 4681−4686.

(27) Peters, G. M.; Davis, J. T. Supramolecular gels made from nucleobase, nucleoside and nucleotide analogs. *Chem. Soc. Rev.* **2016**, *45*, 3188−3206.

(28) Mondal, S.; Das, S.; Nandi, A. K. A review on recent advances in polymer and peptide hydrogels. *Soft Matter* **2020**, *16*, 1404−1454.

(29) Shao, T.; Falcone, N.; Kraatz, H.-B. Supramolecular peptide gels: Influencing properties by metal ion coordination and their wide-ranging applications. *ACS Omega* **2020**, *5*, 1312−1317.

(30) Nebot, V. J.; Smith, D. K. In *Functional Molecular Gels*; Escuder, B., Miravet, J. F., Eds.; Royal Society of Chemistry: London, 2013; pp 30−67.

(31) Wallace, M.; Iggo, J. A.; Adams, D. J. Using solution state NMR spectroscopy to probe NMR invisible gelators. *Soft Matter* **2015**, *11*, 7739−7747.
(32) Onogi, S.; Shigemitsu, H.; Yoshii, T.; Tanida, T.; Ikeda, M.; Kubota, R.; Hamachi, I. In situ real-time imaging of self-sorted supramolecular nanofibres. *Nat. Chem.* **2016**, *8*, 743−752.
(33) Draper, E. R.; Dietrich, B.; McAulay, K.; Brasnett, C.; Abdizadeh, H.; Patmanidis, I.; Marrink, S. J.; Su, H.; Cui, H.; Schweins, R.; Seddon, A.; Adams, D. J. Using small-angle scattering and contrast matching to understand molecular packing in low molecular weight gels. *Matter* **2020**, *2*, 764−778.
(34) Durmus, A.; Gunbas, G.; Farmer, S. C.; Olmstead, M. M.; Mascal, M.; Legese, B.; Cho, J. Y.; Beingessner, R. L.; Yamazaki, T.; Fenniri, H. Synthesis of N-substituted pyrido[4,3-d]pyrimidines for the large-scale production of self-assembled rosettes and nanotubes. *J. Org. Chem.* **2013**, *78*, 11421−11426.
(35) Wang, L.; Partridge, B. E.; Huang, N.; Olsen, J. T.; Sahoo, D.; Zeng, X. B.; Ungar, G.; Graf, R.; Spiess, H. W.; Percec, V. Extraordinary acceleration of cogwheel helical self-organization of dendronized perylene bisimides by the dendron sequence encoding their tertiary structure. *J. Am. Chem. Soc.* **2020**, *142*, 9525−9536.
(36) Roche, C.; Sun, H. J.; Prendergast, M. E.; Leowanawat, P.; Partridge, B. E.; Heiney, P. A.; Araoka, F.; Graf, R.; Spiess, H. W.; Zeng, X. B. O.; Ungar, G.; Percec, V. Homochiral columns constructed by chiral self-sorting during supramolecular helical organization of hat-shaped molecules. *J. Am. Chem. Soc.* **2014**, *136*, 7169−7185.
(37) Peterca, M.; Percec, V.; Imam, M. R.; Leowanawat, P.; Morimitsu, K.; Heiney, P. A. Molecular structure of helical supramolecular dendrimers. *J. Am. Chem. Soc.* **2008**, *130*, 14840−14852.
(38) Chen, L.; Morris, K.; Laybourn, A.; Elias, D.; Hicks, M. R.; Rodger, A.; Serpell, L.; Adams, D. J. Self-assembly mechanism for a naphthalene-dipeptide leading to hydrogelation. *Langmuir* **2010**, *26*, 5232−5242.
(39) Adams, D. J.; Morris, K.; Chen, L.; Serpell, L. C.; Bacsa, J.; Day, G. M. The delicate balance between gelation and crystallisation: structural and computational investigations. *Soft Matter* **2010**, *6*, 4144.
(40) Morris, K. L.; Chen, L.; Raeburn, J.; Sellick, O. R.; Cotanda, P.; Paul, A.; Griffiths, P. C.; King, S. M.; O'Reilly, R. K.; Serpell, L. C.; Adams, D. J. Chemically programmed self-sorting of gelator networks. *Nat. Commun.* **2013**, *4*, 6.
(41) Karunakaran, S. C.; Cafferty, B. J.; Weigert-Munoz, A.; Schuster, G. B.; Hud, N. V. Spontaneous symmetry breaking in the formation of supramolecular polymers: Implications for the origin of biological homochirality. *Angew. Chem., Int. Ed.* **2019**, *58*, 1453−1457.
(42) Kashida, H.; Hattori, Y.; Tazoe, K.; Inoue, T.; Nishikawa, K.; Ishii, K.; Uchiyama, S.; Yamashita, H.; Abe, M.; Kamiya, Y.; Asanuma, H. Bifacial nucleobases for hexaplex formation in aqueous solution. *J. Am. Chem. Soc.* **2018**, *140*, 8456−8462.
(43) Hud, N. V.; Schultze, P.; Sklenar, V.; Feigon, J. Binding sites and dynamics of ammonium ions in a telomere repeat DNA quadruplex. *J. Mol. Biol.* **1999**, *285*, 233−243.
(44) Phillips, K.; Dauter, Z.; Murchie, A. I. H.; Lilley, D. M. J.; Luisi, B. The crystal structure of a parallel-stranded guanine tetraplex at 0.95 angstrom resolution. *J. Mol. Biol.* **1997**, *273*, 171−182.
(45) Menor Salvan, C.; Bouza, M.; Fialho, D. M.; Burcar, B. T.; Fernandez, F. M.; Hud, N. V. Prebiotic origin of pre-RNA building blocks in a urea "warm little pond" scenario. *ChemBioChem* **2020**, *21*, 3504−3510.
(46) Cafferty, B. J.; Fialho, D. M.; Khanam, J.; Krishnamurthy, R.; Hud, N. V. Spontaneous formation and base pairing of plausible prebiotic nucleotides in water. *Nat. Commun.* **2016**, *7*, 11328.
(47) Cafferty, B. J.; Hud, N. V. Was a pyrimidine-pyrimidine base pair the ancestor of Watson-Crick base pairs? Insights from a systematic approach to the origin of RNA. *Isr. J. Chem.* **2015**, *55*, 891−905.
(48) Hud, N. V.; Cafferty, B. J.; Krishnamurthy, R.; Williams, L. D. The origin of RNA and 'My Grandfather's Axe'. *Chem. Biol.* **2013**, *20*, 466−474.
(49) Fialho, D. M.; Clarke, K. C.; Moore, M. K.; Schuster, G. B.; Krishnamurthy, R.; Hud, N. V. Glycosylation of a model proto-RNA nucleobase with non-ribose sugars: implications for the prebiotic synthesis of nucleosides. *Org. Biomol. Chem.* **2018**, *16*, 1263−1271.
(50) Gellert, M.; Lipsett, M. N.; Davies, D. R. Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. U. S. A.* **1962**, *48*, 2013−2018.
(51) Watson, J. D.; Crick, F. H. C. Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid. *Nature* **1953**, *171*, 737−738.
(52) Franklin, R. E.; Gosling, R. G. Evidence for 2-chain helix in crystalline structure of sodium deoxyribonucleate. *Nature* **1953**, *172*, 156−157.
(53) Liu, K. L.; Sasisekharan, V.; Miles, H. T.; Raghunathan, G. Structure of Py•Pu•Py DNA triple helices. Fourier transforms of fiber-type X-ray diffraction of single crystals. *Biopolymers* **1996**, *39*, 573−589.
(54) Patterson, A. L. The Scherrer formula for x-ray particle size determination. *Phys. Rev.* **1939**, *56*, 978−982.
(55) Stubbs, G. J. Effect of disorientation on intensity distribution of non-crystalline fibers. II. Applications. *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr.* **1974**, *30*, 639−645.
(56) Fraser, R. D. B.; Macrae, T. P.; Miller, A.; Rowlands, R. J. Digital processing of fiber diffraction patterns. *J. Appl. Crystallogr.* **1976**, *9*, 81−94.
(57) Chothia, C.; Levitt, M.; Richardson, D. Helix to helix packing in proteins. *J. Mol. Biol.* **1981**, *145*, 215−250.
(58) Parker, T. M.; Hohenstein, E. G.; Parrish, R. M.; Hud, N. V.; Sherrill, C. D. Quantum-mechanical analysis of the energetic contributions to pi stacking in nucleic acids versus rise, twist, and slide. *J. Am. Chem. Soc.* **2013**, *135*, 1306−1316.
(59) Fenniri, H.; Packiarajan, M.; Vidale, K. L.; Sherman, D. M.; Hallenga, K.; Wood, K. V.; Stowell, J. G. Helical rosette nanotubes: Design, self-assembly, and characterization. *J. Am. Chem. Soc.* **2001**, *123*, 3854−3855.
(60) Fenniri, H.; Tikhomirov, G. A.; Brouwer, D. H.; Bouatra, S.; El Bakkari, M.; Yan, Z.; Cho, J.-Y.; Yamazaki, T. High field solid-state nmr spectroscopy investigation of ^{15}N-labeled rosette nanotubes: Hydrogen bond network and channel-bound water. *J. Am. Chem. Soc.* **2016**, *138*, 6115−6118.
(61) Petelski, A. N.; Guerra, C. F. Hydrogen-bonded rosettes of aminotriazines for selective-ion recognition. *J. Phys. Chem. C* **2020**, *124*, 3352−3363.
(62) Engelhart, A. E.; Plavec, J.; Persil, O.; Hud, N. V. In *Nucleic Acid-Metal Ion Interactions*; Hud, N. V., Ed.; RCS Publishing, 2009; pp 118−153.
(63) Peters, J. P.; Maher, L. J. I. In *Bacterial Chromatin: Methods and Protocols, Methods in Molecular Biology*; Dame, R. T., Ed.; Humana Press: New York, NY, 2018; pp 211−256.
(64) Bustamante, C.; Bryant, Z.; Smith, S. B. Ten years of tension: single-molecule DNA mechanics. *Nature* **2003**, *421*, 423−427.
(65) Kawakatsu, T. *Statistical Physics of Polymers: An Introduction*; Springer-Verlag: Heidelberg, 2004.
(66) Peters, J. P.; Mogil, L. S.; McCauley, M. J.; Williams, M. C.; Maher, L. J., III Mechanical properties of base-modified DNA are not strictly determined by base stacking or electrostatic interactions. *Biophys. J.* **2014**, *107*, 448−459.
(67) Kruse, H.; Banas, P.; Sponer, J. Investigations of stacked DNA base-pair steps: Highly accurate stacking interaction energies, energy decomposition, and many-body stacking effects. *J. Chem. Theory Comput.* **2019**, *15*, 95−115.
(68) Peterca, M.; Imam, M. R.; Ahn, C. H.; Balagurusamy, V. S. K.; Wilson, D. A.; Rosen, B. M.; Percec, V. Transfer, amplification, and inversion of helical chirality mediated by concerted interactions of C_3-supramolecular dendrimers. *J. Am. Chem. Soc.* **2011**, *133*, 2311−2328.
(69) Horowitz, E. D.; Engelhart, A. E.; Chen, M. C.; Quarles, K. A.; Smith, M. W.; Lynn, D. G.; Hud, N. V. Intercalation as a means to

https://doi.org/10.1021/jacs.0c12010
J. Am. Chem. Soc. 2021, 143, 6079−6094

suppress cyclization and promote polymerization of base-pairing oligonucleotides in a prebiotic world. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 5288−5293.
(70) Shao, Y.; Gan, Z.; Epifanovsky, E.; Gilbert, A. T. B.; Wormit, M.; Kussmann, J.; Lange, A. W.; Behn, A.; Deng, J.; Feng, X. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184−215 others.
(71) Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **2006**, *27*, 1787−1799.
(72) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
(73) Dunning, T. H., Jr. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(74) Boys, S. F.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* **1970**, *19*, 553−566.
(75) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. Use of approximate integrals in ab initio theory. An application in MP2 energy calculations. *Chem. Phys. Lett.* **1993**, *208*, 359−363.
(76) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796−6806.
(77) Csaszar, A. G.; Allen, W. D.; Schaefer, H. F., III In pursuit of the ab initio limit for conformational energy prototypes. *J. Chem. Phys.* **1998**, *108*, 9751−9764.
(78) East, A. L. L.; Allen, W. D. The heat of formation of NCO. *J. Chem. Phys.* **1993**, *99*, 4638−4650.
(79) Burns, L. A.; Marshall, M. S.; Sherrill, C. D. Comparing counterpoise-corrected, uncorrected, and averaged binding energies for benchmarking noncovalent interactions. *J. Chem. Theory Comput.* **2014**, *10*, 49−57.
(80) Marshall, M. S.; Burns, L. A.; Sherrill, C. D. Basis set convergence of the coupled-cluster correction, δ MP2 CCSD (T): Best practices for benchmarking non-covalent interactions and the attendant revision of the S22, NBC10, HBC6, and HSG databases. *J. Chem. Phys.* **2011**, *135*, 194102.
(81) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. Basis-set convergence in correlated calculations on Ne, N_2, and H_2O. *Chem. Phys. Lett.* **1998**, *286*, 243−252.
(82) DePrince, A. E., III; Sherrill, C. D. Accurate noncovalent interaction energies using truncated basis sets based on frozen natural orbitals. *J. Chem. Theory Comput.* **2013**, *9*, 293−299.
(83) Jeziorski, B.; Moszynski, R.; Szalewicz, K. Perturbation theory approach to intermolecular potential energy surfaces of van der Waals complexes. *Chem. Rev.* **1994**, *94*, 1887−1930.
(84) Papajak, E.; Zheng, J.; Xu, X.; Leverentz, H. R.; Truhlar, D. G. Perspectives on basis sets beautiful: Seasonal plantings of diffuse basis functions. *J. Chem. Theory Comput.* **2011**, *7*, 3027−3034.
(85) Parrish, R. M.; Burns, L. A.; Smith, D. G. A.; Simmonett, A. C.; DePrince, A. E., III; Hohenstein, E. G.; Bozkaya, U.; Sokolov, A. Y.; Di Remigio, R.; Richard, R. M. Psi4 1.1: An open-source electronic structure program emphasizing automation, advanced libraries, and interoperability. *J. Chem. Theory Comput.* **2017**, *13*, 3185−3197.
(86) Zoete, V.; Cuendet, M. A.; Grosdidier, A.; Michielin, O. SwissParam: a fast force field generation tool for small organic molecules. *J. Comput. Chem.* **2011**, *32*, 2359−2368.
(87) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2009**, *31*, 671−690 others.
(88) Dodda, L. S.; Cabeza de Vaca, I.; Tirado-Rives, J.; Jorgensen, W. L. LigParGen web server: an automatic OPLS-AA parameter generator for organic ligands. *Nucleic Acids Res.* **2017**, *45*, W331−W336.
(89) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(90) Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E., III; Cruzeiro, V. W. D.; T.A, D.; R.E, D.; Ghoreishi, D.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; D, G.; Harris, R.; Homeyer, N.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, R. T. a.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; F, P.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. *AMBER* 2018; University of California: San Francisco, 2018.
(91) Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L.-P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comput. Biol.* **2017**, *13*, No. e1005659.
(92) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(93) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(94) Schneider, T.; Stoll, E. Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1978**, *17*, 1302−1322.
(95) McDonald, M.; Kendall, A.; Tanaka, M.; Weissman, J. S.; Stubbs, G. Enclosed chambers for humidity control and sample containment in fiber diffraction. *J. Appl. Crystallogr.* **2008**, *41*, 206−209.
(96) Effenberger, H.; Mereiter, K.; Zemann, J. Crystal-structure refinements of magnesite, calcite, rhodochrosite, siderite, smithonite, and dolomite, with discussion of some aspects of the stereochemistry of calcite type carbonates. *Z. Kristallogr. - Cryst. Mater.* **1981**, *156*, 233−243.
(97) Bian, W.; Wang, H.; McCullough, I.; Stubbs, G. WCEN: a computer program for initial processing of fiber diffraction patterns. *J. Appl. Crystallogr.* **2006**, *39*, 752−756.

https://doi.org/10.1021/jacs.0c12010
J. Am. Chem. Soc. 2021, 143, 6079−6094